VATT Working Papers 112

Assessing the Productivity Consequences of Agri-Environmental Practices When Adoption Is Endogenous

AJ A. Bostian
Moriah B. Bostian
Marita Laukkanen
Antti Simola

VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

112

Assessing the Productivity Consequences of Agri-Environmental Practices When Adoption Is Endogenous

AJ A. Bostian
Moriah B. Bostian
Marita Laukkanen
Antti Simola

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Helsinki 2018

AJ A. Bostian, School of Social Sciences and Humanities, University of Tampere, Finland

Moriah B. Bostian, Department of Economics, Lewis & Clark College, United States

Marita Laukkanen, VATT Institute for Economic Research and Department of Economics, University of Helsinki, Finland

Antti Simola, VATT Institute for Economic Research, Finland

We are grateful to several colleagues for their helpful comments: Shawna Grosskopf, David Zirkle; seminar participants at Portland State University, University of Helsinki, Umeå University Centre for Environmental and Resource Economics (CERE), VATT Institute for Economic Research, University of Alabama, and University of Sydney; and conference participants at APPC, EAERE, and EAAE. This work was partially supported by grants from Lewis & Clark College and the William J. Fulbright Foundation.

ISBN 978-952-274-225-4 (PDF)

ISSN 1798-0291 (PDF)

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, December 2018

Assessing the Productivity Consequences of Agri-Environmental Practices When Adoption Is Endogenous

VATT Institute for Economic Research
VATT Working Papers 112/2018

AJ A. Bostian – Moriah B. Bostian – Marita Laukkanen – Antti Simola

This Version: November 18, 2018

Abstract

We address the general problem of selection bias, endemic to analyzing the effects of any policy where adoption is voluntary, with empirical application to environmental policies for agriculture. Many voluntary practices for mitigating the environmental impacts of agriculture provide external benefits while lowering productivity. Policy analysis of the productivity consequences is complicated by the fact that decision-makers can choose their own policy levers, an action that ruins any notion of random assignment. We introduce an identification strategy to correct this kind of endogeneity, combining classic methods from stochastic frontier analysis and selection models. Applying it to micro-level data from Finnish grain farms, we find that more efficient producers are more likely to enroll in subsidized practices. And, because these practices tend to reduce yield, frontier analysis without the endogeneity correction greatly understates productivity losses. In other words, naively basing the frontier estimator on the subset of less productive farms leads to downward bias in the resulting frontier estimates. In fact, average inefficiency more than doubles after the correction in this case. An outlier investigation also suggests that the lowest decile of farms are responsible for most of the selection bias in the uncorrected model.

Key words: productivity, stochastic frontier analysis, endogeneity, selection model, agrienvironmental policy

JEL classes: Q53, Q58, Q18, Q12, D24, C54, C34, C36

1 Introduction

Agriculture is a significant source of non-point pollution in many countries. Because these contaminants are difficult to regulate centrally (e.g., due to asymmetric information), mitigation often relies on decentralized measures. In effect, farms are often handed a set of pre-approved policy levers, and told to make the best possible outcome for themselves. Agri-environmental programs (AEPs), policies that subsidize voluntary adoption of environmental practices, are an example of this approach.

Because many AEPs are inherently yield-reducing (e.g., fertilizer restrictions), they present ambiguous welfare implications: environmental improvements can be offset by productivity losses. To see the full welfare picture, policymakers need accurate estimates of both facets. But, the fact that farms choose their own policy levers greatly complicates any assessment of the productivity part. Choice from a menu patently violates bedrock random-assignment assumptions used in empirical productivity analysis.

A farm facing a set of policy options will adopt only the ones that satisfy a cost-benefit criterion particular to that farm. Consequently, AEP policy analysis should begin by investigating this adoption step. If adoption is instead taken as given, the policy's efficiency impact will be biased. To illustrate, consider an AEP menu with two options. A farm will evaluate not only the overt incentives attached to each option (e.g., the subsidy amount per area), but also how well each one integrates with the farm's infrastructure (e.g., the substitutability of crops across specific land tracts). Different farms can rationally come to different conclusions on the integration question, and choose differently (and non-randomly) on that basis.

Suppose that a large fraction of farms adopt one practice – what does that say about the productivity of the other? A naïve answer would simply note the first's output dominance, and declare the second inefficient on that basis. A more nuanced answer would wrap the farm's integration rationale into the output assessment, recognizing that farms adopting the second policy do so from a reasoned appraisal of their own costs and benefits. In this light, the selection of an uncommon practice is not an inherently inefficient choice.

To accommodate this and other selection stories, we develop a two-stage empirical model of the productivity consequences of AEP adoption. The production stage is rooted in stochastic frontier analysis (SFA) (Aigner et al., 1977), a method for estimating a best-practices frontier of outputs for given inputs. SFA's principal identifying restriction decomposes the residual total factor productivity into explicitly nonnegative inefficiencies, and all other unobservable productivity effects. Somewhat unusually, the quantity to be identified is not a parameter, but the residual decomposition itself. Because moments of the inefficiency term feed into widely-used efficiency indexes, measurements of this residual have important practical ramifications (Kumbhakar and Lovell, 2000).

The adoption stage is rooted in classic selection models (Heckman, 1979). Framed as a selection problem, AEP endogeneity arises because a farm assesses the potential inefficiencies across the menu when deciding which practices to adopt. Although the possibility of inefficiency-based endogeneity is raised frequently in the productivity literature, most of the endogeneity discussion to date has focused on simultaneity in production inputs. Amsler et al. (2016) compare the performance of various empirical strategies used to correct this kind of endogeneity. Horrace et al. (2016) apply a polychotomous, Heckman-style correction to a network production model that assigns workers to groups. We are aware of only one other study that addresses endogenous technology choice in a policy context: Kumbhakar et al. (2009) jointly estimate a dairy farm's decision to go organic, and the productivity consequences of those organic processes.

Our identification strategy enhances this last kind of endogeneity correction in two ways. First, by using a multinomial framework, we can model more than two options in the adoption stage, thereby addressing a broader spectrum of policies. Second, we allow for arbitrary correlations among the inefficiency terms. This captures the fact that a farm's potential inefficiencies are probably grounded in common sources. For example, an AEP menu of only fertilizer restrictions implicates similar farm processes involving fertilizer.

To our knowledge, ours is the first empirical study to account for endogeneity of more than two decision options within an SFA-style framework, while simultaneously allowing arbitrary inefficiency correlations. Our application has particular significance for policy evaluation of AEPs in the US and EU, the areas with highest uptake of such programs. But, this endogeneity correction is not limited to AEPs. It applies in generic situations where a decision-maker evaluates a menu of technological options before proceeding to production.

We estimate the model on a repeated cross-section of Finnish grain farms observed during Finland's third AEP policy period, 2007-2014. We find that correcting for adoption endogeneity drastically alters the naïve efficiency assessment: among farms adopting an AEP, average inefficiency more than doubles after the correction. Moreover, *efficient* farms are more likely to select into AEPs.

Because the Finnish AEPs all tend to reduce yield, the latter finding carries worrisome welfare implications. However, it may be driven by highly unproductive outliers. We find that the relationship between uptake and efficiency fully inverts after removing the farms in the lowest efficiency decile, a result more compatible with the likely policy objectives.

Because the model is highly nonlinear, we conduct a brief simulation study for intuition about its asymptotics. A toy calibration indicates that the Central Limit Theorem (CLT) applies only when sample sizes reach 10^3 to 10^4 observations. Unfortunately, even when perfectly specified, models estimated on 10^2 observations tend to have flat sampling distributions, particularly in the adoption stage. Thus, employing this method requires substantial amounts of data. Fortunately, our application has about 5,000 observations, making us relatively confident about the asymptotics in our case.

2 Policy Background

The EU's Common Agricultural Policy (CAP) establishes EU-wide environmental objectives, but leaves policy design largely up to member states. The EU has historically emphasized working-land AEPs, while the US has enacted them only recently (Baylis et al., 2008). These practices often require major alterations to farm processes, raising the specter of productivity losses. As AEPs expand in use, quantifying these productivity consequences grows increasingly important.

Finland provides an ideal case study of that tradeoff. More than 90% of Finnish agricultural land is enrolled in some kind of AEP, compared to 25% across the fifteen older EU members (European Commission, 2005). Finland also enrolls the most land – about 95% in Finland versus 26% across the EU – in AEPs specifically requiring input reduction. These practices entail sizeable productivity impacts.

Finland enacted its AEP policy upon joining the EU in 1995. Unlike the programs of most member states, its program applies nationwide. Because the policy's scope has expanded significantly since its inception, we focus on Finland's third AEP period, 2007-2014. To ensure that we model a coherent technological space, we investigate only the active farms that cultivated grain during that time.[1]

Finland's policy has two participation levels: "basic measures" and "additional and/or special measures." For grain farms, basic measures for reducing nutrient loading impose fertilizer limits by crop and region, and require 3 m filter strips for crops adjacent to waterways. The "additional measures" and "special measures" (which require concurrent enrollment in basic measures) include further reductions in fertilizer input, more accurate application of fertilizer, reduced tillage, winter cover crops, long-term crop rotation, nutrient balance, riparian zones, and organic processes. Subsidies vary by measure and crop, and participation ranges from 10-30% for the additional and special measures.[2] This variation in participation rates, coupled with farm-level detail on output, allows us to empirically decompose the productivity consequences of adopting these AEPs.

This critical, first-order relationship has been largely unexplored in the literature. As we noted previously, we know of only one other study (Kumbhakar et al., 2009) that has actually corrected for inefficiency-based endogeneity, and only in the context of binary uptake.[3] The authors find that more inefficient dairy farms are less likely to adopt organic practices, even though organic subsidies are available. Moreover, those

[1] In Finland, "active farms cultivating grain" include both crop farms with at least 1 ha under cultivation, and animal farms with at least 1 ha under cultivation or at least one animal being raised.

[2] The "Rural Development Programme for Mainland Finland 2007-2013" describes the eligibility criteria, subsidy amounts, and participation rates. Basic-measure subsidies for crop farms are 93 EUR/ha, and participation exceeds 90%. The non-binding character of the fertilizer limits may explain this very high uptake rate (Laukkanen and Nauges, 2014).

[3] Some policy analysis of French agriculture has been conducted in this vein in the context of input endogeneity, but outside the SFA literature (Piot-Lepetit and Moing, 2007; Mary, 2013).

farms going organic are ultimately about 5% less efficient than those that do not, an appreciable welfare tradeoff.

A few policy studies examine the efficiency of US and Canadian AEPs (Zhao et al., 2004; Feng et al., 2006; Rabotyagov et al., 2010; Tamini et al., 2012). Tamini et al. (2012) investigate the use of manure injection and herbicide reduction in Canadian crop farms. Applying a dataset with farm and environmental outcomes, they estimate the "environmental efficiency" of a farm's decisions, along with productivity and profit. They find that herbicide reductions decrease technical efficiency, but that the freed labor and capital also increase productivity. Manure injection slightly increases technical efficiency and productivity, and substantially increases profitability. And, with regard to our core welfare concern, the authors find a positive relationship between technical and environmental efficiencies.

The literature on EU AEPs focuses almost exclusively on general efficacy and environmental outcomes (Ohl et al., 2008; Wätzold et al., 2008; Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2013; Laukkanen and Nauges, 2014). Laukkanen and Nauges (2014) estimate input use exploiting regional differences in Finland's AEP payments, and simulate the resulting effects on nutrient loadings. They find that basic water-protection measures, such as vegetative filter strips and winter cover crops, reduce nutrient loadings by about 10%. Other aspects previously explored include farms' participation decision (Hynes et al., 2008; Giovanopoulou et al., 2011) and the spatial distribution of measures (Wätzold and Drechsler, 2005; Bamière et al., 2011). We contribute to this stream of research by estimating the productivity consequences of three of Finland's most intensive practices.

3 Correcting for Adoption Endogeneity

We introduce the model by describing the interlinked adoption and production decisions. Each of I farms first chooses from a menu of J AEP practices, and then produces output using the selected practice. Starting with the production step and working backward, farm

i's output under practice j is given by the log-linear Cobb-Douglas production model

$$y_{ij} = x'_{ij}\beta_j + \varepsilon_{ij}$$

where y_{ij} is (log) output, x_{ij} is a (log) vector of observable input factors, and ε_{ij} are unobservable input factors. The unobservables are further decomposed into explicit inefficiencies u_{ij} (positive by definition) and all other production characteristics v_{ij}, so that $\varepsilon_{ij} = -u_{ij} + v_{ij}$. This residual structure mirrors classic SFA.

The classic framework also assumes random assignment of practices to farms. As we have already argued, nothing like random assignment happens in reality. A farm instead chooses optimally from the menu of practices specified in the AEP regulation, based on the manager's assessment of how profitable each practice is likely to be. Farm i's selection of practice j is driven by the expected-profit model

$$\pi_{ij} = z'_i\gamma_j + \delta u_{ij} + \eta_{ij}$$

where z_i is a vector of observed profitability factors, and η_{ij} are unobserved profitability factors. From our perspective as researchers, we observe the AEP choice j and the resulting output y_{ij}, but not the manager's set of J intermediate expected-profit calculations π_{ij}. These latent judgments, however, are the key drivers of the two outcomes we can observe: uptake and output.

So far, this setting is much like a classic two-stage selection model. The fly in the ointment is the presence of efficiency considerations in the selection step. Naïvely ignoring this endogeneity would incorrectly attribute efficiencies to the policy menu that actually arise from the farm's choice of practices that best suit its characteristics.

To frame the issue in likelihood terms, suppose we observe farm i implementing AEP practice $m_i = j$ and producing $y_i = y_{ij}$. Denote the vector of J inefficiencies that farm i incurs under the menu of AEP policies as u_i. From our perspective as researchers, the likelihood $f(y_i, m_i)$ of this occurring is

$$
\begin{aligned}
f(y_i, m_i = j) &= E_{u_i}\left[f(y_i, m_i = j | u_i)\right] \\
&= E_{u_i}\left[f(y_i | m_i = j, u_i) f(m_i = j | u_i)\right] \\
&= E_{u_i}\left[f(v_{ij} | m_i = j, u_i) f(m_i = j | u_i)\right] \\
&= \int_{\mathbb{R}^{J+}} f(v_{ij} | m_i = j, u_i) f(m_i = j | u_i)\, dF(u_i)
\end{aligned}
$$

This likelihood provides correct productivity inferences in the presence of inefficiency-based endogeneity.

When likelihoods like this one entail a complex relationship among unobservables, the empirical literature often makes simplifying assumptions to maintain tractability. Unfortunately, we cannot make any of the usual ones without abandoning key features of the AEP setting. One common simplification is that m_i and v_{ij} are independent, i.e., that adoption is random with respect to non-inefficiency unobservables. This assumption is effectively random assignment, and collapses the model back to SFA. But, in reality, the unobserved factors driving output for farm i under practice j are almost assuredly similar to the ones that led the farm to adopt it in the first place.

Another standard simplification is that the inefficiencies are all independent, an assumption that conveniently reduces $dF(u_i)$ to a product of likelihoods. However, an AEP policy menu typically implicates common farm operations, and so the elements of u_i are likely to be correlated instead. Indeed, one of the practices in the Finnish policy menu is a combination of two others, guaranteeing correlation almost automatically.

We can maintain tractability while preserving the requisite correlations by placing additional structure on the unobservables. In the output equation, we implement a common SFA-style normal/half-normal form for ε_{ij}. That is, the conditional density of the production unobservables v_{ij} is normal

$$
v_{ij} | m_i = j, v_{ij} \; \sim \; N\left(0, \sigma_{vj}^2\right)
$$

and the unconditional density of inefficiencies u_i is half-normal

$$u_i \sim N^+ (0, \Sigma_u)$$

Importantly, unlike standard SFA, the distribution of u_i has an unrestricted covariance matrix Σ_u.[4]

To aid in interpreting estimates, it is worthwhile to note the units of the $f(v_{ij}|m_i = j, u_i)$ and $f(u_i)$ distributions. The units of the production equation are the units of y_{ij}: log t of grain output in our case. As a result, $x'_{it}\beta$, u_{it}, and v_{it} also have units of log t. The conditional value $v_{ij}|m_i = j, u_i$ can be written as a difference in logs

$$v_{ij} = y_{ij} - \left(x'_{ij}\beta - u_{ij}\right)$$

and so small values of v_{ij} can also be interpreted as farm i's percentage distance from frontier j. Indeed, a common SFA interpretation of $f(v_{ij}|m_i = j, u_i)$ is a "distribution of percent deviations." However, the same interpretation is not valid for u_{ij} here, because we do not analogously perform the reverse conditioning. That is, we are intentionally *not* modeling $f(u_{ij}|m_i = j, v_{ij})$, which would provide the same percent-distance interpretation for small values of u_{ij}. We hence have nothing to say about that distribution. We are instead modeling $f(u_{ij})$ unconditionally, and so its units are simply log t of output.[5] In fact, because log t is always a valid unit of comparison in this context, we discuss all productivity consequences in those terms.

In the selection equation, we use a standard Gumbel distribution for the unobserved profitability factors η_{ij}:

$$\eta_{ij}|u_i \sim GEV\left(0, \sigma^2_\eta, 1\right)$$

[4]Standard multivariate SFA arises when $\Sigma_u^{1/2}$ is diagonal.

[5]Unconditional u_{ij} also provides more useful input to the selection equation. Because adoption is based on profit levels, other level variables are appropriate predictors for that equation, rather than percent changes that have no scale.

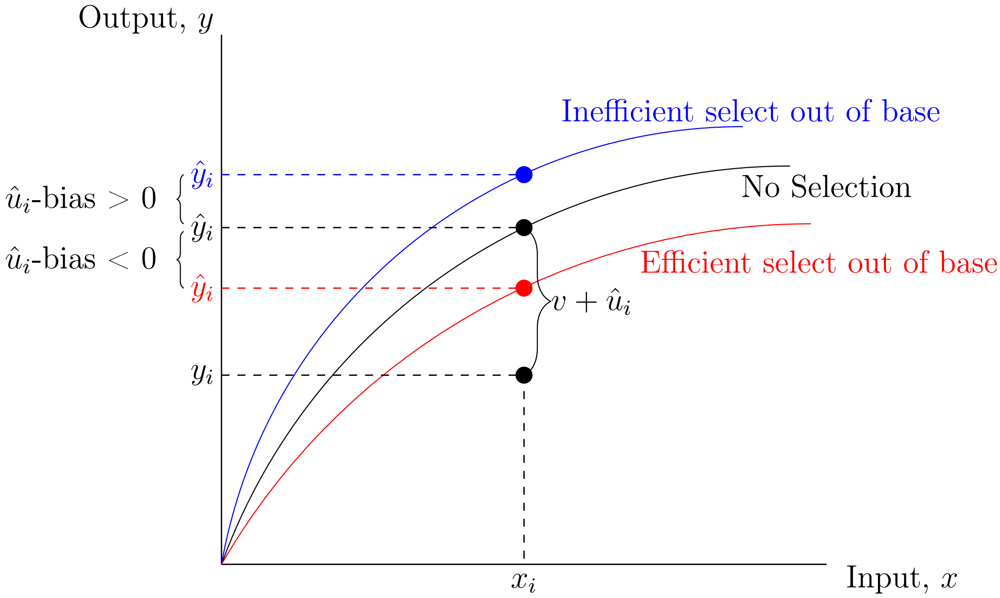


Figure 1: Stochastic frontier with endogeneity (base group case)

This yields a logit probability for the adoption of practice j:

$$f\left(m_i = j | u_i\right) = \frac{\exp\left(z_i' \gamma_j + \delta u_{ij}\right)}{1 + \sum_{j'=1}^{J-1} \exp\left(z_i' \gamma_{j'} + \delta u_{ij'}\right)}$$

We normalize the γ_j parameters relative to the Jth practice, and the shape parameter σ_η^2 to 1.

With first-stage logit probabilities inserted, the likelihood takes the slightly friendlier form

$$f\left(y_i, m_i = j\right) = \int_{\mathbb{R}^{J+}} \left[\frac{1}{\sigma_{vj}} \phi\left(\frac{y_i - x_{ij}' \beta_j + u_{ij}}{\sigma_{vj}}\right)\right] \left[\frac{\exp\left(z_i' \gamma_j + \delta u_{ij}\right)}{1 + \sum_{j'=1}^{J-1} \exp\left(z_i' \gamma_{j'} + \delta u_{ij'}\right)}\right] dF\left(u_i\right)$$

But, this expression is still much more complex than the SFA likelihood, which is actually closed-form:

$$f\left(y_i, m_i = j\right) = \left[\frac{1}{\sigma_{vj}} \phi\left(\frac{y_i - x_{ij}' \beta_j + u_{ij}}{\sigma_{vj}}\right)\right] \cdot \left[\frac{1}{\sigma_{uj}} \frac{\phi\left(u_{ij}/\sigma_{uj}\right)}{\Phi\left(u_{ij}/\sigma_{uj}\right)}\right]$$

Comparing the two reveals a new tractability issue stemming from u_i's flexible covariance. Namely, the integral over $\mathbb{R}^{J+}$ has potentially high dimension ($J = 4$ in our case), and such integrals are difficult to evaluate numerically. This is a nuisance problem for the policy analysis, but failing to address it results in estimates that are overly contaminated with approximation error. As we note Section 4, this endogeneity correction requires a fairly large sample size to overcome its small-sample bias, and we do not need to return hard-won gains in statistical accuracy by injecting a major source of numerical bias. We discuss two integration algorithms in Appendix A, and implement the one that exhibits less error.

With regard to identification, the most important quantities are the first two moments of u_{ij}, because those attributes feed into standard efficiency indexes. We thus forego causal interpretations of β, γ, and δ here. If those interpretations are needed in other applications, it is certainly possible to extend this method to incorporate them.

For similar reasons, the potential decision-theoretic downsides of the selection specification (e.g., the independence of irrelevant alternatives) do not particularly concern us. Again, we make no attempt here to causally interpret γ or δ, or the tradeoffs they imply. For our purposes, the only role of the selection stage is to carve out an exogenous slice of the total variation in adoption decisions, one that can re-condition the production stage into something approximating random assignment. The success of this undertaking depends primarily on the selection equation's *nonparametric* features (e.g., strong and exogenous instruments), not its parametric ones (e.g., logit vs. probit structure) (Matzkin, 2007).

4 Finite-Sample Behavior

Because the model above is highly nonlinear, we conduct a brief simulation study to get intuition about its finite-sample properties. Of course, when the model is correctly specified and the sampling process is random, the CLT guarantees that the maximum-likelihood estimators are consistent and asymptotically normal. But, the requirement for employing the CLT ($I \to \infty$) provides little guidance on whether it actually holds in our context. We thus investigate the behavior of a toy calibration in sample sizes frequently encountered in empirical work: $I = 100$ and $I = 1000$ observations.

This example uses $J = 2$ options, a single-input production stage, and a singleinstrument profit stage. Hence, x and z are 2×1 vectors (including a constant in each). The vectors in an (x, z) pair are drawn disjointly, making the calibration nonparametrically identified by construction. The error variances are set so that about half of the variation in each equation comes from the observables, and half from the unobservables. Finally, the elements of u_i are uncorrelated. These would be very favorable circumstances to find in a real application, in our view.

Table 1: Means, standard deviations, and standard skews of the simulated sampling distribution.

		$I = 100$			$I = 1000$		
	True	Mean	Std. Dev.	Std. Skew	Mean	Std. Dev.	Std. Skew
β_{11}	0.25	0.24	0.33	0.30	0.21	0.15	-0.07
β_{12}	1.00	1.00	0.13	-0.01	1.00	0.04	-0.02
σ_{v1}	0.38	0.34	0.13	-0.31	0.39	0.05	-0.19
β_{21}	0.50	0.52	0.21	0.14	0.49	0.07	0.10
β_{22}	0.75	0.75	0.11	-0.13	0.75	0.03	0.03
σ_{v2}	0.22	0.18	0.08	0.22	0.22	0.03	0.03
γ_{21}	0.19	0.80	8.31	1.91	0.19	1.31	0.87
γ_{22}	-0.19	-1.00	3.67	-5.23	-0.26	0.22	-3.36
δ	0.80	8.90	90.74	5.42	0.64	7.50	4.07
σ_{u11}	0.45	0.64	0.50	2.44	0.68	0.32	1.56
σ_{u22}	0.32	0.40	0.23	0.79	0.40	0.14	1.59
σ_{u12}	0.00	-0.42	0.67	-1.67	-0.34	0.49	-1.93

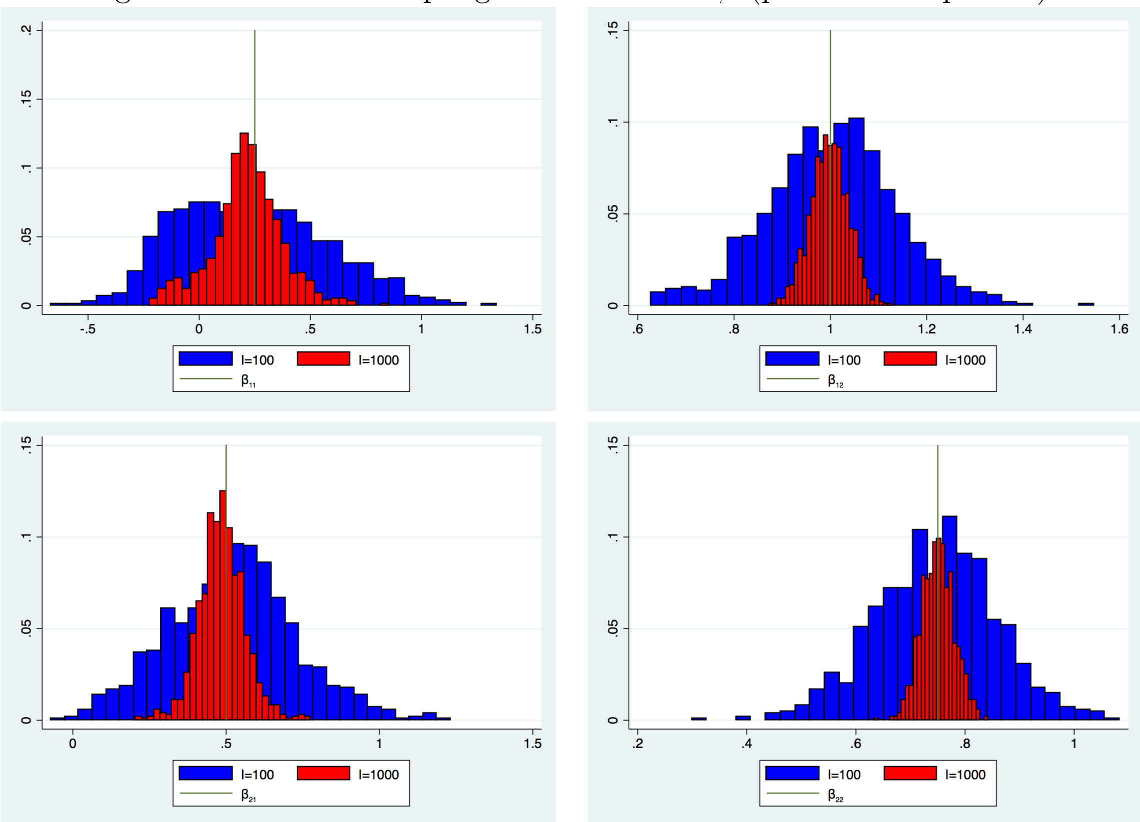
Table 1 compares the true parameter values to the first three moments of their simulated sampling distributions. For each value of I, we simulated one thousand datasets. The second-stage β estimators appear to be centered fairly well at the correct value, as the CLT would predict. However, their standard deviations are quite large when $I = 100$, a finding corroborated the spread of their histograms in Figure 2. These estimators are thus marginally trustworthy when $I = 100$, but fairly robust when $I = 1000$.

The first-stage γ and δ estimators behave quite poorly when $I = 100$, indicating that the CLT does not yet apply. The center of γ is off by a factor of four, and the standard deviation is multiple times the center.[6] Overall, convergence is much slower in the selection equation: the center moves in the correct direction by $I = 1000$, but the standard deviation is still large. The sampling distribution of δ, which is almost

[6]The $I = 100$ histograms in Figure 3 are actually truncated to the range of the $I = 1000$ histograms, because a meaningful amount of the $I = 100$ density lies more than five standard deviations from the mean.

completely flat when $I = 100$, starts to become more defined by $I = 1000$. Even so, it is still about 20% too low, and retains very high spread. The γ estimators are thus reasonably robust when $I = 1000$, but the δ estimator is still not very reliable.

Figure 2: Simulated sampling distributions of β (production equation).



The σ_v estimators appear to be trustworthy by $I = 100$ at first glance. However, their nearly-correct central values belie a double peak in their sampling distributions, a feature evident only from the histograms in Figure 4. Of course, the CLT permits no bimodality, and so these estimators are actually quite poor. However, they are once again reliable by $I = 1000$.

The Σ_u estimators converge slowly, and frequently uncover spurious correlation between the inefficiencies. The histogram of the off-diagonal covariance element in Figure 5 reveals a correct modal estimate near 0, accompanied by a long left tail. This skew persists even by $I = 1000$, a feature at odds with the symmetric distribution predicted by the CLT. Thus, this estimator is questionable even by $I = 1000$.

To summarize, β, γ, and σ_v are reasonably well-behaved by $I = 1000$, but δ and Σ_u are not. This finding is unfortunate, because addressing the endogeneity of u_i requires reliable estimates of Σ_u and δ. This slow convergence occurs in nearly ideal conditions: the toy calibration has no specification error, single-variable equations, guaranteed nonparametric identification, and about half of its outcome variation explained by observables.

Figure 3: Simulated sampling distributions of γ and δ (selection equation).

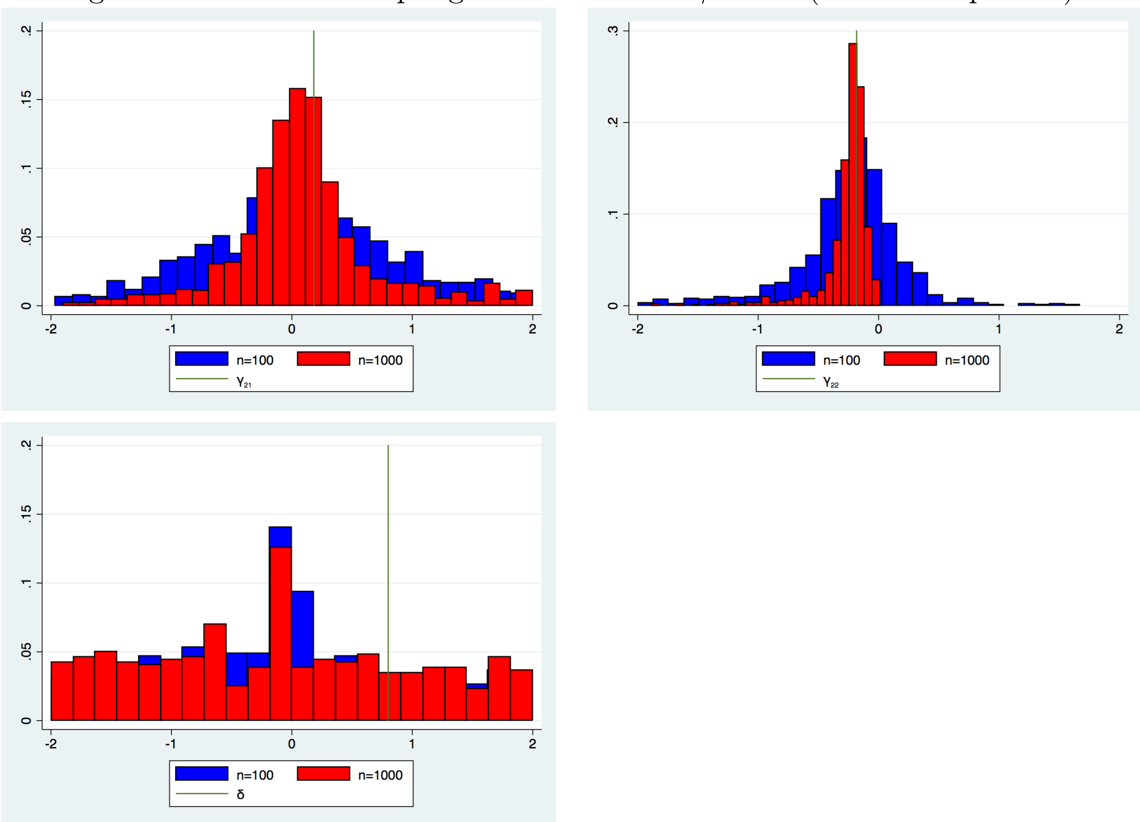
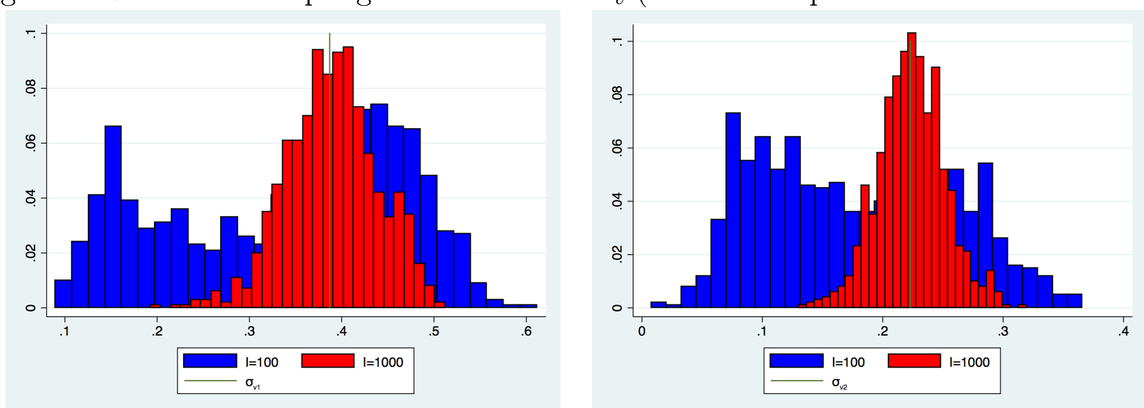
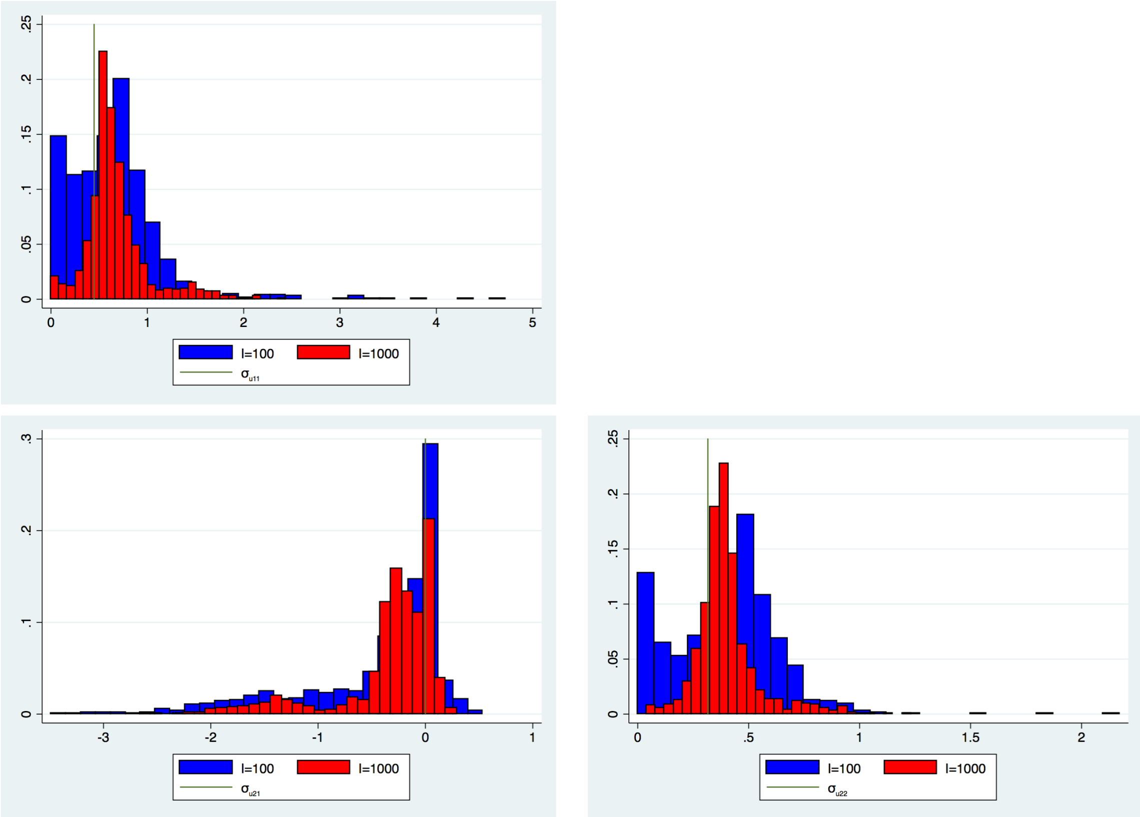


Figure 4: Simulated sampling distributions of σ_v (unobserved production characteristics).



The practicalities of the CLT thus appear front and center when applying this endogeneity correction. The method is very much a large-sample one: even if the model is perfectly specified, estimates with samples of only hundreds are probably unreliable. Our application contains about $I = 5000$ observations and generates mostly precise estimates, giving us more confidence about its asymptotics.

Figure 5: Simulated sampling distributions of σ_u (unobserved inefficiencies).



5 Data

Our micro-level farm data come from two Finnish registry sources. The first is the database “Statistics on the Finances of Agricultural and Forestry Enterprises” maintained by Statistics Finland. It consists of a detailed yearly input/output survey at the farm level, as well as tax-registry records on incomes, expenses, assets, and liabilities. The survey is a rolling sample of farms interviewed for two consecutive years, making our final dataset a repeated cross-section.

Production variables are almost always recorded as revenues and expenditures. We back out the underlying quantities by referring to price indexes from the “Yearbook of Farm Statistics” and the “Producer Prices of Agricultural Products,” both published by

Table 2: “Additional measures” enacted during second and third AEP periods.

Policy Number	Second Period (2000-2006)	Third Period (2007-2014)
1	More accurate fertilization	Reduced fertilization
2	-	More accurate nitrogen fertilization on arable crops
3	Plant cover in winter and reduced tillage	Plant cover in winter and reduced tillage
4	Biodiversity on farms	Plant cover in winter
5	Reduction in ammonia emissions	Intensified plant cover in winter
6	Improving the welfare of animals	Crop diversification
7	Treatment of washing water from milking rooms	Extensive grassland production
8	More accurate follow up of nutrients	Extensive grassland production, region C
9	Measuring soluble nitrogen	Spreading of manure during the growing season
10	Organic cover in weed control	Nutrient balance
11	Catchment of manure gases	Cultivation of catch plants
12	-	More accurate nitrogen fertilization of horticultural crops
13	-	Use of mulch in perennial horticultural crops
14	-	Use of pest monitoring methods

Note: The official policy numbers of some practices changed between AEP periods.

the Information Centre of the Ministry of Agriculture and Forestry.

AEP participation is recorded in the database “Farm Accounting and Income Statistics” maintained by Finland’s Agency for Rural Affairs. This identifies each farm’s land enrollments, and the subsidies paid.

Finland finances its AEPs jointly with the EU. The first policy period (1995-1999), which immediately followed Finland’s entry into the EU, largely consisted of flat region- and crop-specific basic payments for a few practices. It was subsequently revised in conjunction with the Agenda 2000 CAP reform. The resulting second period (2000-2006) introduced more flexibility through a wider menu of “additional” and “special” measures. In 2003, the Fischler reform fundamentally changed the CAP, ushering in a third policy period (2007-2014). This rather substantial amendment increased Finland’s AEP budget, and again altered the menu of eligible practices.

The “Statistics on the Finances of Agricultural and Forestry Enterprises” database has partial coverage of the second AEP period, and full coverage of the third. Table 2 lists the "additional measures" in effect during each. Although some policy overlap exists, many new measures were introduced in the third period, and some of the second-period ones were refined or discarded. The sheer number of changes leads us to treat the third period as a policy break, and so we investigate it in isolation. With this focus, we can estimate efficiency relationships without developing an ancillary control strategy for differences in policy implementation.

A few statistics illuminate the main differences between the two periods. Figure 6 presents the time series of payments to basic, additional, and special measures. Finland has always subsidized more than EUR 200 M of basic measures, but a persistent jump in subsides of additional, and to a lesser extent, special measures occurred at the start of the third period. Figures 7 and 8 list the average land areas enrolled in each practice, by period. The upsurge in payments is reflected in the land enrollments as well, with widespread adoption of accurate nitrogen application (practice 2) and winter cover crops / reduced tillage (practice 3). Of course, farms were not limited to one practice; Figures 9 and 10 present the same period averages by combination of practices. Practices 2 and 3 were most often used, either in isolation or in conjunction with other practices.

Figure 6: AEP spending, second and third AEP periods.

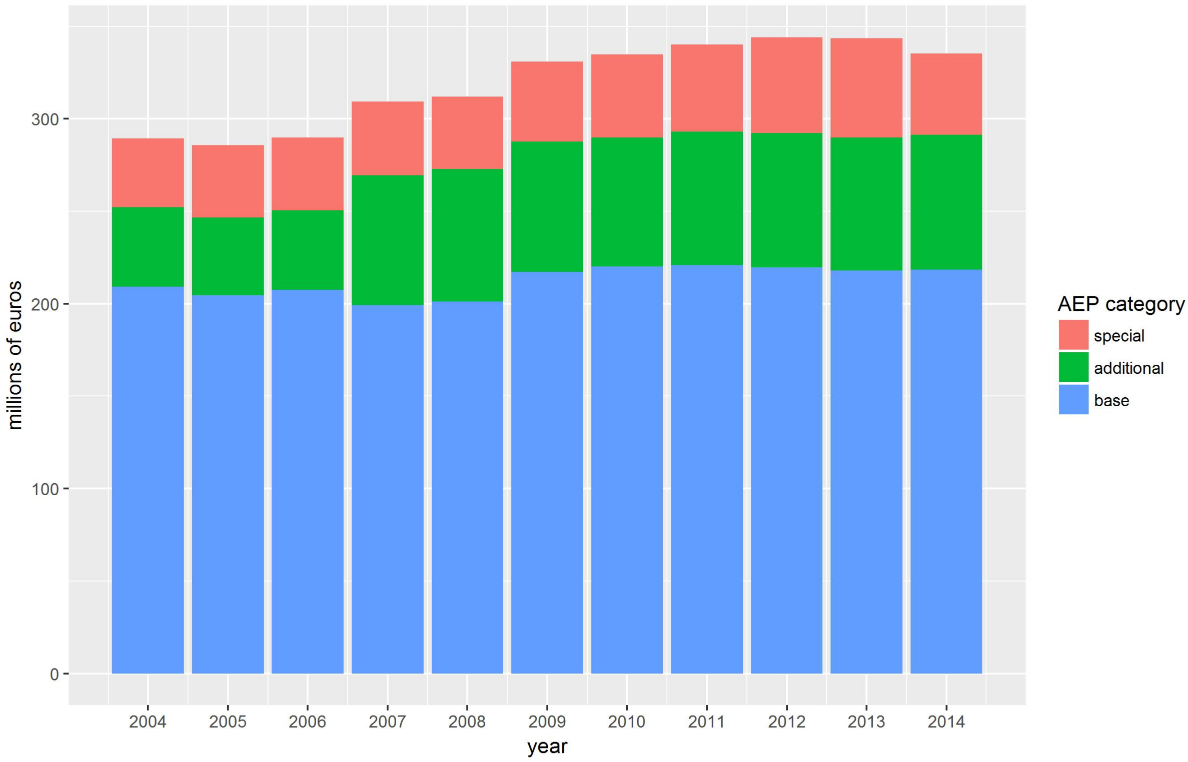


Figure 7: Average annual land area enrolled in "additional measures," second AEP period.

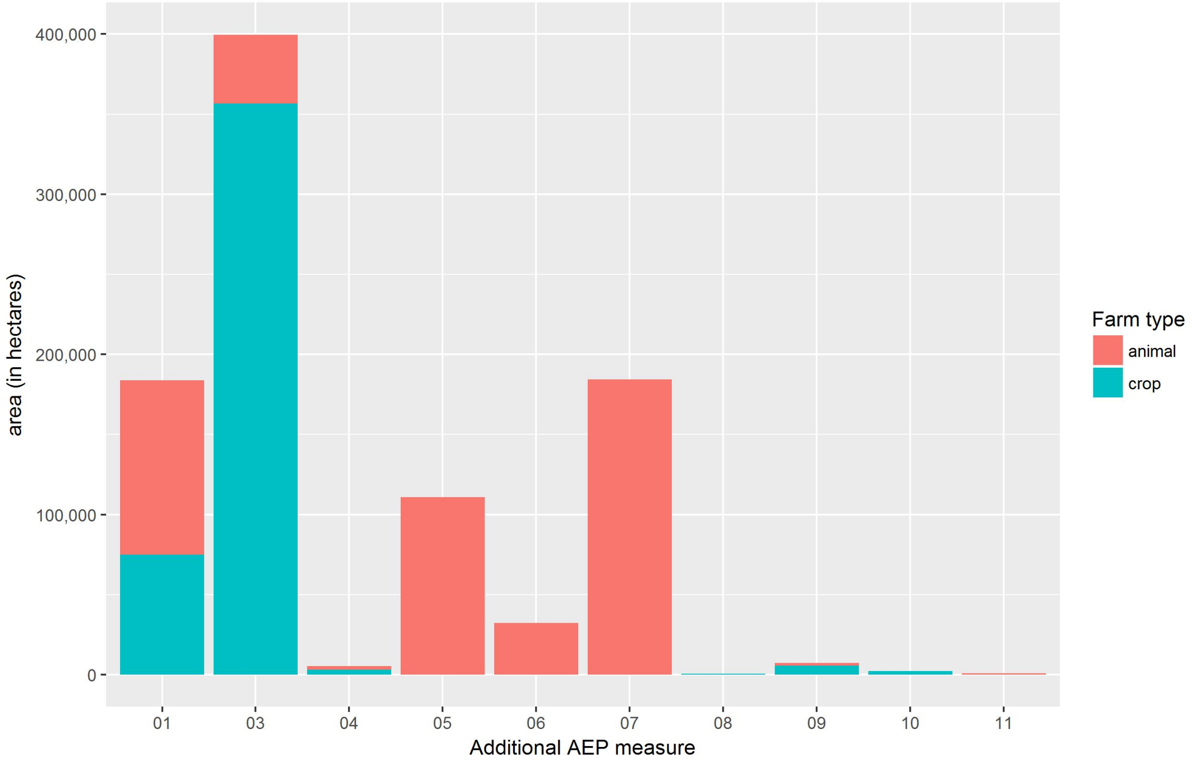


Figure 8: Average annual land area enrolled in "additional measures," third AEP period.

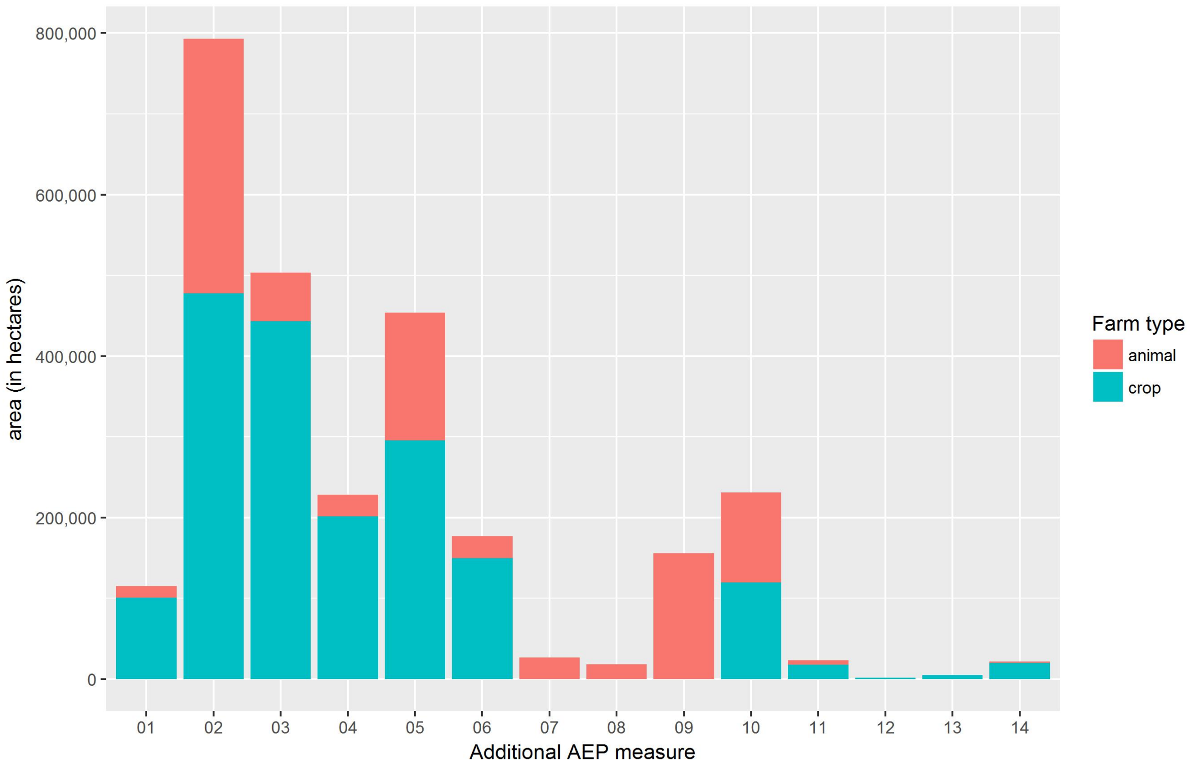


Figure 9: Average annual land area enrolled in "additional measure" combinations, second AEP period.

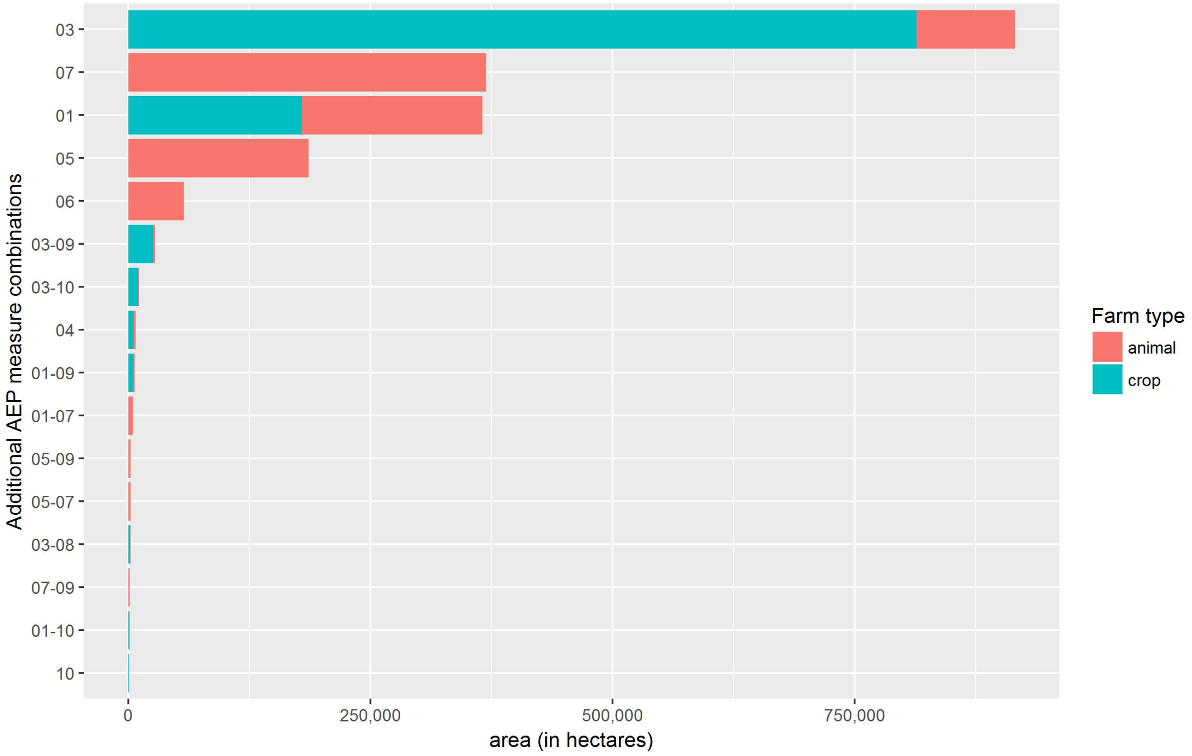
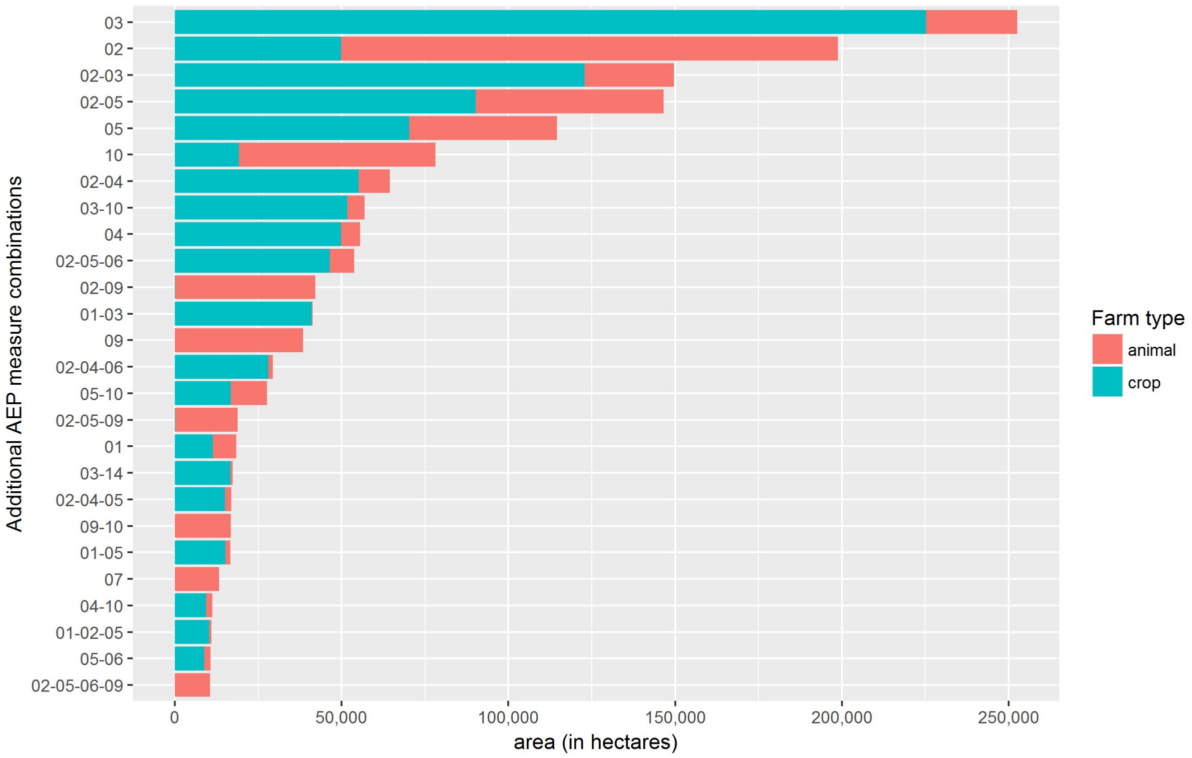


Figure 10: Average annual land area enrolled in "additional measure" combinations, third AEP period.



Because these two practices are adopted at a much higher rate than the others, we investigate an AEP menu of practice 2 alone, practice 3 alone, and the combined 2 & 3 practice. This leaves us with enough farms in each group to generate sufficient variation in adoption. We compare them all to a baseline group that did not enroll in any additional measures, one third of which did not enroll in even the basic program. We thus evaluate $J = 4$ options, indexing them by popularity of the practice: $j = 1$ for practice 3, $j = 2$ for practice 2, $j = 3$ for practice 2 & 3, and $j = 4$ for baseline non-adopters.

Figure 11: Finnish subsidy regions within the study area.

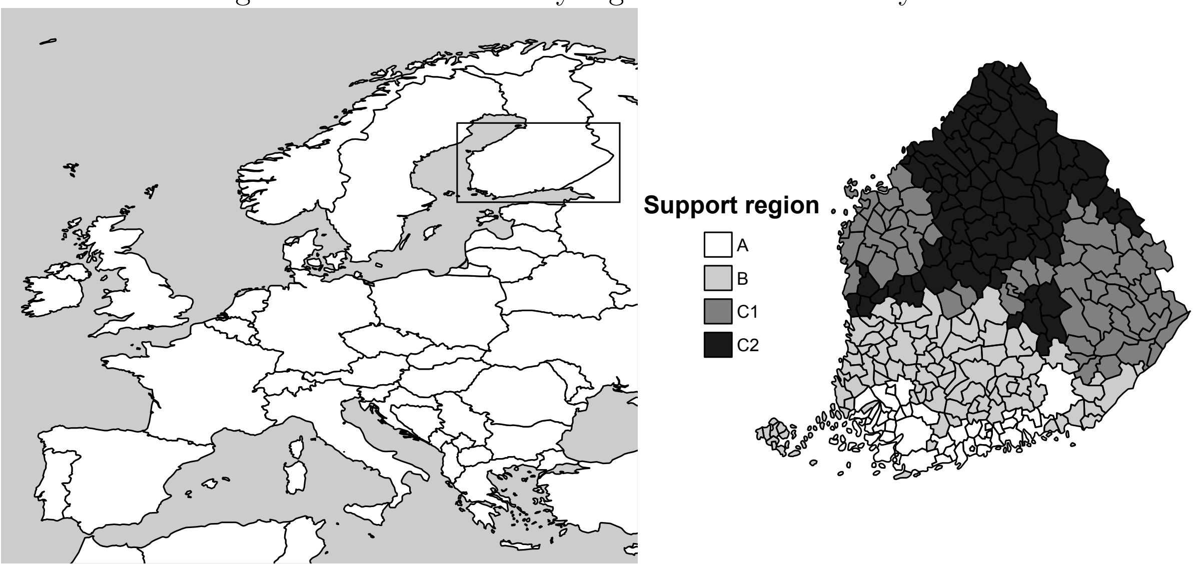


Table 3 summarizes the data.[7] The production equation includes the primary inputs land, fertilizer, machinery, and fuel.[8] The remaining variables are geographic and temporal frontier shifters. Finnish agricultural subsidies vary over seven "support regions" drawn to local climates. We focus on the four regions A-C2 (see Figure 11), an area containing 98% of Finland's grain production.[9] Because the Baltic Sea can induce microclimates, we include an indicator of whether the farm sits near it. We capture additional climate and weather variability with average temperature, precipitation, and growing-season length.

[7]Complete registry data for 2014 had not been released at the time of this analysis.

[8]Labor is omitted because most Finnish farms do not employ workers outside the family.

[9]Farms in the other three regions are located in the far north (partly above the Arctic Circle), and face substantially different growing conditions than rest of the country.

Table 3: Summary statistics.

	Practice 3 (1,893 obs.)		Practice 2 (1,615 obs.)		Practice 2 & 3 (739 obs.)		Not Adopting (446 obs.)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Grain (t)	147.99	160.17	79.44	78.75	145.65	150.05	59.18	52.41
Land (ha)	86.58	72.10	44.25	36.20	73.93	57.64	29.73	22.08
Fertilizer (t)	71.54	87.49	33.82	42.55	59.93	63.65	24.38	25.44
Machinery (index)	419.53	619.78	206.74	290.12	377.18	538.10	143.10	212.42
Fuel (hl)	60.78	35.59	31.86	38.03	57.91	68.33	21.07	26.12
2007	0.16	0.37	0.20	0.40	0.18	0.38	0.22	0.41
2008	0.15	0.36	0.15	0.36	0.13	0.33	0.17	0.38
2009	0.15	0.36	0.14	0.35	0.14	0.35	0.15	0.36
2010	0.14	0.34	0.14	0.34	0.16	0.37	0.10	0.30
2011	0.13	0.34	0.13	0.33	0.16	0.36	0.13	0.34
2012	0.14	0.34	0.13	0.34	0.11	0.31	0.10	0.30
2013	0.14	0.34	0.11	0.31	0.13	0.33	0.12	0.33
Region A	0.36	0.48	0.13	0.34	0.21	0.41	0.03	0.16
Region B	0.55	0.50	0.15	0.36	0.25	0.43	0.17	0.37
Region C1	0.08	0.27	0.41	0.49	0.28	0.45	0.56	0.50
Region C2	0.02	0.12	0.30	0.46	0.27	0.44	0.25	0.43
Temperature (avg C)	14.01	0.98	13.66	0.93	13.88	0.94	13.59	0.91
Precipitation (avg cm)	2.16	0.45	2.29	0.48	2.22	0.44	2.29	0.48
Growing Season (days)	153.57	11.74	145.78	15.42	146.93	15.54	144.86	14.75
Baltic	0.67	0.47	0.73	0.44	0.73	0.44	0.71	0.45
Manager's Age (years)	51.49	11.71	51.69	12.41	50.55	11.93	52.54	12.16
CAP Subsidy (EUR 1000)	18.901	16.983	8.530	6.958	14.808	12.512	5.667	4.157
LFA Subsidy (EUR 1000)	16.968	13.812	9.251	7.537	15.136	11.593	6.107	5.302
Debts/Assets	1.04	7.82	0.68	1.34	0.75	1.13	0.52	1.21
Municipal Practice 3 (avg rate)	0.57	0.17	0.28	0.21	0.48	0.20	0.27	0.17
Municipal Practice 2 (avg rate)	0.22	0.17	0.63	0.27	0.55	0.27	0.49	0.24
Municipal No Adoption (avg rate)	0.03	0.07	0.10	0.11	0.08	0.10	0.25	0.17

The selection equation includes instruments for selection bias. First, we include attributes of the farm and its management. We proxy managerial experience with the manager's age. We proxy financial health with a farm's debt-to-asset ratio, and the subsidy payments it receives under the EU CAP and Less-Favoured Area (LFA) programs. We conjecture that financially-secure farms are more likely adopt AEPs because they can better absorb any unforeseen profit shock associated with the process changes.

We also include average adoption rates at the municipality level. Because municipalities are much smaller than support regions, these averages capture a shared local rationale for choosing specific practices. As a result, they correlate reasonably well with farm-level adoption decisions. At the same time, municipal-level adoption rates probably do not enter an individual farm's output decision. Hence, they are candidates for strong and exogenous instruments.

Farms adopting practice 3 tend to be larger, more input-intensive, and more highly subsidized through CAP and LFA. They are also more likely to be found in the southern regions A and B. The size and input use of farms adopting practice 2 generally fall between those adopting practice 3 and the non-adopters. These farms are more concentrated in the more northern regions C1 and C2. Like those adopting practice 3, adopters of the combined practice tend to be larger farms, but they are more evenly distributed across support regions. Finally, non-adopters are found primarily in regions B-C2, and are concentrated most highly in C1. They receive lower EU subsidies (likely due to their smaller size), are less indebted, and are more likely to be found in municipalities with higher rates of non-adoption.

6 Results

To highlight the importance of the endogeneity correction, we begin with reduced-form estimates of each decision stage: log-linear and SFA models of production, and multinomial logit models of adoption. We follow the reduced form results with the full set of endogeneity-corrected second-stage results, and finally consider these in light of the fist stage.

Table 4: Pooled log-linear and SFA estimates.

	Log-Linear		SFA	
	(I)	(II)	(I)	(II)
Constant	**10.893**	**11.337**	**7.413**	**7.719**
Land	**0.825**	**0.842**	**0.887**	**0.895**
Fertilizer	**0.096**	**0.094**	**0.073**	**0.071**
Machinery	-0.005	-0.003	0.001	0.001
Fuel	*0.021*	*0.021*	**0.025**	**0.025**
2008	**-0.164**	**-0.171**	-0.061	*-0.065*
2009	-0.015	-0.013	0.039	0.040
2010	*-0.087*	*-0.079*	-0.057	-0.052
2011	0.069	*0.077*	0.050	0.050
2012	**-0.103**	**-0.099**	**-0.077**	**-0.083**
2013	**-0.299**	**-0.288**	**-0.201**	**-0.186**
Region B	**0.160**	**0.163**	**0.044**	*0.037*
Region C1	-0.028	**-0.089**	**-0.102**	**-0.139**
Region C2	**-0.101**	**-0.174**	**-0.173**	**-0.219**
Temperature	**-2.366**	**-2.439**	**-1.435**	**-1.496**
Precipitation	-0.067	-0.064	-0.007	0.002
Growing Season	**-0.778**	**-0.816**	**-0.473**	**-0.490**
Baltic	**0.092**	**0.087**	**0.098**	**0.092**
Practice 3		**-0.177**		**-0.115**
Practice 2		**-0.079**		**-0.078**
Practice 2 & 3		-0.046		0.006
σ_v			**0.205**	**0.202**
σ_u			**0.925**	**0.924**
R^2	0.609	0.611		

Bold: significant at 5%. *Italics:* significant at 10%.

In the pooled log-linear estimates in Table 4, each of the primary inputs has the expected sign (except machinery, which is insignificant). Unsurprisingly, these estimates reveal that Finnish grain production is land- and fertilizer-intensive. More curiously, temperature and growing-season duration both factor negatively. Nonlinearities in the marginal effect of weather could be responsible for this result (Schlenker and Roberts, 2008). Relative to base region A, output is higher in region B and lower in regions C1 and C2. Also, farms near the Baltic Sea tend to have higher output. Relative to base year 2007, output tends to be lower over the rest of the policy period.

Table 5: Log-linear (i), SFA (ii), and endogeneity-corrected (iii) estimates by policy group.

	Practice 3			Practice 2			Practice 2 & 3			Not Adopting		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Constant	**14.696**	**9.507**	**14.710**	**9.699**	**8.120**	**9.712**	**9.426**	4.153	**9.442**	-4.788	-2.425	-4.771
Land	**0.846**	**0.891**	**0.880**	**0.826**	**0.898**	**0.874**	**0.915**	**0.898**	**0.907**	**0.810**	**0.875**	**0.809**
Fertilizer	**0.095**	**0.074**	**0.084**	**0.104**	**0.074**	**0.077**	**0.082**	**0.074**	**0.080**	0.055	**0.054**	*0.055*
Machinery	-0.020	-0.012	-0.009	0.012	*0.014*	*0.015*	0.015	0.008	0.008	-0.008	**0.039**	-0.007
Fuel	0.030	**0.040**	**0.037**	0.005	0.009	0.011	0.014	*0.037*	0.025	0.021	0.003	0.018
2008	*-0.178*	-0.069	**-0.188**	**-0.188**	*-0.113*	**-0.151**	-0.109	0.040	-0.076	0.152	*0.176*	0.166
2009	-0.045	0.034	-0.004	0.031	0.018	0.008	-0.099	-0.046	-0.051	0.087	*0.169*	0.015
2010	*-0.168*	**-0.167**	-0.079	0.000	0.039	0.069	-0.139	-0.112	-0.053	-0.235	-0.104	-0.203
2011	0.033	0.019	0.097	0.059	0.053	0.079	0.125	0.061	*0.191*	-0.063	-0.013	-0.061
2012	**-0.156**	**-0.146**	**-0.179**	-0.084	*-0.093*	**-0.111**	-0.005	0.071	0.064	0.180	0.120	0.179
2013	-0.145	-0.107	-0.019	**-0.334**	**-0.200**	**-0.179**	**-0.558**	**-0.452**	**-0.355**	**-0.836**	**-0.438**	**-0.756**
Region B	**0.082**	-0.010	0.002	**0.205**	0.048	0.038	**0.302**	**0.153**	*0.141*	**0.466**	*0.221*	**0.478**
Region C1	-0.014	-0.016	-0.103	-0.084	**-0.196**	**-0.215**	-0.081	-0.117	*-0.172*	**0.490**	0.159	**0.492**
Region C2	**-0.308**	**-0.270**	**-0.332**	**-0.167**	**-0.260**	**-0.279**	-0.082	-0.126	*-0.215*	**0.488**	0.131	**0.474**
Temperature	**-3.099**	**-1.917**	**-3.083**	**-2.094**	**-1.546**	**-2.008**	-1.765	-0.474	**-1.720**	1.562	0.565	1.621
Precipitation	-0.103	-0.063	-0.099	0.043	0.054	0.050	*-0.329*	-0.223	**-0.301**	-0.004	0.132	0.041
Growing Season	**-1.140**	**-0.620**	**-1.067**	**-0.713**	**-0.573**	**-0.650**	**-0.829**	-0.314	**-0.736**	0.238	0.376	0.277
Baltic	0.014	0.033	0.040	**0.148**	**0.104**	**0.077**	0.099	**0.155**	**0.127**	**0.221**	**0.141**	**0.221**
σ_v		**0.235**	**0.310**		**0.161**	**0.206**		**0.218**	**0.308**		**0.120**	**0.450**
σ_u		**0.963**	**2.013**		**0.846**	**2.587**		**0.953**	**2.458**		**0.843**	**1.112**
R^2	0.540			0.603			0.590			0.553		

Bold: significant at 5%. *Italics:* significant at 10%.

Including practice indicators does not materially change any of the pooled results. But, the large coefficients on these indicators confirm that AEP choices are indeed correlated with output, even after controlling for typical inputs. This is a sure sign of adoption endogeneity.

The SFA estimates in Table 4 largely mirror the log-linear results. The notable exceptions are the coefficients on weather and growing-season duration, both of which shrink by about 40%. The time indicators also uniformly shrink (up to 37%), and there are large but directionally unsystematic changes in the region indicators (decreases up to 73%, increases up to 264%). SFA's identifying restriction thus reclassifies many regional, time, and weather effects as outright inefficiency. And, those inefficiencies are relatively large: the estimated standard deviation of u_i is 0.558 log t. A simple inefficiency index, the mean inefficiency $E(u_i)$, can be constructed from this estimate. In this case, inefficiency averages 0.738 log t, or about 2% of pooled mean output.

Moving to the endogeneity correction, Table 5 compares the second-stage production estimates to their reduced-form analogs by policy group. This correction does not alter the primary input coefficients within groups very much. The group-level estimates do reveal a large, positive, and statistically significant effect for Baltic Sea proximity, except for farms adopting practice 3. Weather and growing-season effects also substantially expand in magnitude, and there are large but unsystematic differences among the time and region indicators.

Table 6: Mean inefficiency $E(u_i)$ in log t and as a percentage of mean output.

	(II)		(ii)		(iii)	
Practice 3	0.738	1%	0.768	1%	1.606	3%
Practice 2	0.738	3%	0.675	2%	2.064	10%
Practice 2 & 3	0.738	1%	0.760	1%	1.961	5%
Not Adopting	0.738	4%	0.673	3%	0.887	4%

The major differences between the reduced-form and endogeneity-corrected estimates occur in the inefficiency decomposition. Table 6 collects the values of $E(u_i)$ implied by the pooled SFA estimates (II), the uncorrected SFA estimates (ii), and the corrected estimates (iii). The endogeneity correction uncovers substantially higher inefficiencies across the AEP menu. As a percentage of group-level mean output, inefficiency increases from 1% to 3% for practice 3, 2% to 10% for practice 2, 1% to 5% for the combined practice, and 3% to 4% for non-adopters. Thus, endogeneity-corrected inefficiencies for specific AEPs are three to five times higher than their uncorrected values. A similar increase occurs in the baseline group, though not as extreme.

Table 7: Endogeneity-corrected $\Sigma_u^{1/2}$ estimates. Correlation coefficients in parentheses.

	u_{i1}	u_{i2}	u_{i3}	u_{i4}
u_{i1}	**2.013**			
u_{i2}	**-0.585 (-0.036)**	**2.587**		
u_{i3}	-0.544 (-0.104)	-1.704 (-0.328)	**2.458**	
u_{i4}	-0.434 (-0.196)	-0.752 (-0.235)	**-1.541 (-0.531)**	**1.112**

Bold: significant at 5%. *Italics:* significant at 10%.

Table 7 presents the other critical aspect of the inefficiency decomposition, estimates of $\Sigma_u^{1/2}$.[10] These inefficiencies are highly correlated. The off-diagonal estimates tend to be imprecise, but underscore the common operational issues that farms face when choosing from this AEP menu. The bottom row of $\Sigma_u^{1/2}$ shows that as opting out becomes more efficient, there are even greater inefficiencies associated with selecting any AEP. This is consistent with profit-maximizing behavior by non-adopters. A similar relationship occurs when comparing practice 3 to practice 2.

We now turn to the first stage. For maximum statistical efficiency in generating selection variation for the second stage, the multinomial logit includes all variables from the production equation as well as the set of instruments.[11] Because our main concern lies with the instruments, we focus on their exogeneity and strength.[12]

With regard to strength, the reduced-form estimates in Table 8 indicate that farm and managerial characteristics are largely unimportant to the adoption decision, with the exception of the debt/asset ratio for adopters of practice 2 and the combined practice. The municipal average adoption rates, on the other hand, are quite important across all groups, and have the expected signs. For instance, farms in municipalities with higher

[10]For easy comparison between the corrected and uncorrected estimates of $\Sigma_u^{1/2}$, the bottom row of Table 5 contains just its diagonal elements.

[11]Angrist and Pischke (2008, Section 4.6) provide a lucid rationale for this strategy, ultimately concluding, "If a covariate is good enough for the second stage, it's good enough for the first."

[12]Appendix B contains the full set of first-stage estimates.

adoption of practice 2 are less likely to adopt practice 3, as are those located in areas with higher rates of non-adoption. Similarly, farms in municipalities with higher adoption of practice 2 are themselves more likely to adopt practice 2. Finally, farms situated in areas that adopt each practice separately are more likely to choose the combined practice.

Table 8: Multinomial logit (a) and endogeneity-corrected (iii) adoption instruments (point estimates and marginal probability effects at the mean).

	Practice 3		Practice 2		Practice 2 & 3	
	(a)	(iii)	(a)	(iii)	(a)	(iii)
Manager's Age	0.001	0.002	-0.001	0.000	-0.004	-0.004
10 year change	1.2%	1.6%	0.1%	0.1%	-2.1%	-2.2%
Debts/Assets	**0.244**	**0.243**	*0.089*	*0.089*	*0.098*	0.096
1% change	**0.1%**	**0.1%**	*0.0%*	*0.0%*	0.0%	0.0%
CAP Subsidy	-0.017	-0.015	**-0.086**	**-0.085**	-0.032	-0.032
EUR 1000 change	0.3%	0.2%	**-0.6%**	**-0.3%**	-0.9%	-1.0%
LFA Subsidy	0.062	0.061	**0.113**	**0.113**	*0.082*	*0.082*
EUR 1000 change	0.4%	0.7%	**0.4%**	**0.3%**	*2.4%*	*2.4%*
Municipal Practice 3	**2.934**	**2.937**	**-1.632**	**-1.672**	**4.399**	**4.415**
1% change	**0.3%**	**0.4%**	**-0.5%**	**-0.3%**	**1.7%**	**1.5%**
Municipal Practice 2	**-3.789**	**-3.884**	**1.758**	**1.801**	**3.243**	**3.314**
1% change	**-2.1%**	**-2.3%**	**0.1%**	**0.1%**	**2.3%**	**2.2%**
Municipal No Adoption	**-8.752**	**-8.907**	**-6.939**	**-7.303**	**-2.578**	**-2.636**
1% change	**-2.6%**	**-3.0%**	**-0.2%**	**-0.1%**	**0.7%**	**-0.6%**
u_{ij}		**-0.550**		**-0.550**		**-0.550**
1 log t change		**-11.0%**		**-0.4%**		**-13.7%**

Base practice is non-adopters. MNL pseudo-$R^2 = 0.35$.
Bold: significant at 5%. *Italics:* significant at 10%.

Formally evaluating strength is difficult in this setting because the adoption outcome is polychotomous, and there are no F-statistic rules of thumb for this case (Stock et al., 2002). But, we can examine the logit's goodness-of-fit more informally. The pseudo-R^2 of the reduced form is 0.35 with the instruments and 0.24 without, and most of that increase comes from the municipal adoption rates. Thus, the first stage likely contributes a moderate amount of exogenous selection variation to the second stage.

To evaluate exogeneity, we include the candidate instruments in the reduced-form production model, and test for their joint significance (estimates not reported). Except for CAP and LFA payments, they are not jointly significant. This lingering correlation reflects some production attributes that cannot be captured with input, geographic, or temporal controls, but can be proxied with subsidy payments.

With the endogeneity correction in place, we again observe a positive effect of indebtedness in adopting practice 2. LFA subsidies now have a positive effect across the menu, as do CAP subsidies in practice 2 and the combined practice. Very importantly, the instruments responsible for the most exogenous variation, municipal uptake rates, are again large and statistically significant across groups.

The endogeneity-corrected first stage reveals yet another important driver of adoption: the farm's inefficiencies u_i. The strongly negative estimate of δ indicates that efficient farms are much more likely to adopt AEPs. Because these AEPs tend to reduce yield, this selection effect explains the substantially higher inefficiencies that occur with the correction in place. The overall outcome echoes Kumbhakar et al.'s finding that efficient Finnish dairy farms are more likely to adopt organic practices, but that upon doing so, they become less efficient than their non-organic peers.

7 Outliers

The finding that AEP adoption is associated with better efficiency *ex ante* is somewhat disconcerting from a policymaker's perspective. Because these three practices almost assuredly reduce output, the welfare benefits of the AEP policy could be severely curtailed if the more efficient farms are always the ones voluntarily restricting their yields through their choice of levers.

In evaluating this result, it is important to remember that δ reflects an average propensity not only across the menu of options, but also across farms. Consequently, its estimate could be swayed by outliers. In our context, a small number of highly unproductive farms could underpin most of the negative effect. One example is a "hobby farm," an enterprise that does not serve as a main income source but operates more like a recreational project. It meets the "active farm" criteria for tax and AEP purposes, but the manager's objective is not profit maximization in the usual sense.

Our data, though highly disaggregated, are not rich enough to tease out this kind of operational intent. We investigate the possibility more coarsely by excluding outliers. We conjecture that the problematic farms are the ones substantially undershooting the frontier. We thus truncate the dataset using a median absolute deviation (MAD) criterion on the reduced-form fitted values of ε_{ij} (Hampel, 1974):

$$\text{MAD}_{ij} = 1.483 \cdot \text{median} \left(\left| \hat{\varepsilon}_{ij} - \text{median} \left(\hat{\varepsilon_j} \right) \right| \right)$$

Intuitively, an observation with a large MAD value fits the frontier spectacularly badly. But, an important nuance arises when applying MAD in an efficiency context: because ε_{ij} contains both u_{ij} and v_{ij}, a large MAD value is not enough on its own to classify a farm as inefficient. That distinction is an empirical question that our model can address, however.

Our outlier criteria are four (MAD4) and three (MAD3) times the median error.[13] The MAD4 and MAD3 criteria remove successively more farms: 4.4% of them under MAD4 (reductions of 4.3%, 4.3%, 4.7%, and 4.3% in each group), and 8.9% under MAD3 (reductions of 10.2%, 7.3%, 10.1%, and 7.0% in each group).

Table 9 presents the new data means under these outlier criteria. A few characteristics of the high-MAD farms are important. First, mean grain output rises nearly uniformly as the criterion tightens: increases of 3.2%, 3.7%, 4.4%, and 3.5% occur in each group under MAD4; and 3.0%, 3.3%, 4.8%, and 5.1% occur under MAD3. This indicates that the outliers are predominantly lower-output farms. At the same time, mean input usage almost uniformly falls, although the changes are less pronounced. Land use changes by -0.2%, -0.5%, 0.5%, and -2.6% in each group under MAD4; and -1.6%, -1.3%, 0.0%, and -2.0% under MAD3. Fertilizer use changes by -0.6%, 0.2%, 0.4%, -3.5% under MAD4; and -2.9%, -0.1%, 0.1%, and -2.8% under MAD3. The simultaneous rise in mean output and fall in mean inputs strongly suggests that our outliers are small, inefficient farms.

[13] If ε_{ij} were distributed normally, these cutoffs would asymptotically correspond to 4 and 3 standard deviations from the mean.

Table 9: Data means under MAD4 and MAD3.

	Practice 3			Practice 2			Practice 2 & 3			Not Adopting		
	Full	MAD4	MAD3	Full	MAD4	MAD3	Full	MAD4	MAD3	Full	MAD4	MAD3
Observations	1,893	1,812	1,698	1,615	1,547	1,500	739	704	664	446	427	414
Grain (t)	147.99	152.71	151.98	79.44	82.41	82.05	145.65	152.02	152.57	59.18	61.27	62.20
Land (ha)	86.58	86.71	85.21	44.25	44.05	43.67	73.93	74.31	73.96	29.73	28.95	29.13
Fertilizer (t)	71.54	71.14	69.47	33.82	33.91	33.79	59.93	60.17	60.01	24.38	23.52	23.70
Machinery (index)	419.53	422.93	408.24	206.74	207.21	207.72	377.18	380.49	369.16	143.10	141.41	141.66
Fuel (hl)	60.78	60.55	59.46	31.86	31.91	31.67	57.91	58.22	56.22	21.07	20.34	20.23
2007	0.16	0.16	0.17	0.20	0.20	0.20	0.18	0.18	0.18	0.22	0.22	0.22
2008	0.15	0.15	0.15	0.15	0.15	0.15	0.13	0.13	0.13	0.17	0.17	0.17
2009	0.15	0.15	0.15	0.14	0.14	0.15	0.14	0.16	0.14	0.15	0.15	0.15
2010	0.14	0.13	0.13	0.14	0.14	0.14	0.16	0.16	0.16	0.10	0.10	0.10
2011	0.13	0.13	0.13	0.13	0.13	0.12	0.16	0.16	0.16	0.13	0.14	0.14
2012	0.14	0.14	0.14	0.13	0.13	0.13	0.11	0.11	0.11	0.10	0.10	0.10
2013	0.14	0.13	0.13	0.11	0.11	0.10	0.13	0.12	0.11	0.12	0.11	0.11
Region A	0.36	0.35	0.34	0.13	0.13	0.13	0.21	0.21	0.20	0.03	0.02	0.02
Region B	0.55	0.56	0.57	0.15	0.15	0.15	0.25	0.25	0.26	0.17	0.17	0.16
Region C1	0.08	0.08	0.08	0.41	0.41	0.41	0.28	0.27	0.27	0.56	0.56	0.56
Region C2	0.02	0.01	0.02	0.30	0.30	0.31	0.27	0.27	0.27	0.25	0.26	0.26
Temperature (avg C)	14.01	13.99	13.97	13.66	13.65	13.64	13.88	13.87	13.85	13.59	13.58	13.57
Precipitation (avg cm)	2.16	2.16	2.17	2.29	2.29	2.29	2.22	2.22	2.23	2.29	2.30	2.30
Growing Season (days)	153.57	153.55	153.55	145.78	145.78	145.75	146.93	146.91	146.81	144.86	144.85	144.85
Baltic	0.67	0.66	0.66	0.73	0.73	0.73	0.73	0.73	0.73	0.71	0.71	0.72
Manager's Age (years)	51.49	51.48	51.68	51.69	51.72	51.72	50.55	50.68	50.80	52.54	52.52	52.40
CAP Subsidy (1000 EUR)	18.901	18.866	18.402	8.530	8.485	8.389	14.808	14.870	14.691	5.667	5.564	5.583
LFA Subsidy (1000 EUR)	16.968	17.008	16.795	9.251	9.210	9.146	15.136	15.188	15.164	6.107	6.040	6.057
Debts/Assets	1.04	0.80	0.79	0.68	0.69	0.68	0.75	0.74	0.75	0.52	0.51	0.50
Municipal Practice 3 (avg rate)	0.57	0.57	0.57	0.28	0.28	0.28	0.48	0.48	0.48	0.27	0.27	0.26
Municipal Practice 2 (avg rate)	0.22	0.22	0.21	0.63	0.63	0.63	0.55	0.54	0.55	0.49	0.49	0.49
Municipal No Adoption (avg rate)	0.03	0.03	0.03	0.10	0.10	0.10	0.11	0.08	0.08	0.25	0.25	0.25

Most of the other means in Table 9 are robust to MAD tightening, with the exception of EU subsidy payments. Mean CAP payments change by -1.2%, -0.5%, 0.4%, and -1.8% in each group under MAD4; and -1.0%, -1.7%, -0.8%, and -1.5% under MAD3. Mean LFA payments change by 0.2%, -0.4%, 0.3%, and -1.1% under MAD3; and -1.0%, -1.1%, 0.2%, and -0.8% under MAD4. Thus, in addition to being small and inefficient, outlier farms also receive higher subsidies from related EU programs.

Table 10 presents the new first-stage estimates. Interestingly, the estimate of δ rises as the outlier criterion becomes more restrictive. Inefficiency has a small and insignificant negative effect on adoption under MAD4, and a large and significant positive effect under MAD3. To reiterate, the MAD3 estimate of δ is equal in magnitude to the full-sample estimate, but *opposite in sign.* The more efficient farms now appear to be *foregoing* yield-reducing AEPs, a much better welfare outcome.

As Table 11 shows, this inversion of δ also has implications for the second stage. Most critically, σ_u and σ_v fall almost uniformly. This result is not facially surprising, because the MAD procedure eliminates outliers rather than central observations. But, the magnitude of this decline is quite remarkable.

Table 12 collects the implied values of $E\left(u_i\right)$. As a percentage of (full-sample) group-level mean output, inefficiency falls from 3% to 1% for practice 3, 10% to 2% for practice 2, 5% to 1% for the combined practice, and 3% to 2% for non-adopters. These are much more favorable inefficiency levels. Moreover, most of this change occurs during the first trimming of the sample with MAD4, the lightest touch to the data.

Some primary input estimates are also affected by removing these outliers. For example, land's elasticity falls by nearly 0.07 for practice 3, and rises by 0.08 for non-adopters. Temperature and growing-season effects tend to shrink toward 0. Large realignments also occur in the time and geographic indicators, particularly for farms adopting practice 2.

Table 10: Adoption instruments under MAD4 and MAD3 (point estimates and marginal probability effects at the mean).

	Practice 3			Practice 2			Practice 2 & 3		
	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3
Manager's Age	0.002	0.000	0.001	0.000	-0.002	-0.001	-0.004	-0.004	-0.003
10 year change	1.6%	0.8%	1.0%	0.1%	0.0%	0.0%	-2.2%	-1.9%	-1.6%
Debts/Assets	**0.243**	**0.253**	**0.257**	*0.089*	*0.102*	*0.102*	0.096	*0.102*	*0.114*
1% change	**0.1%**	**0.1%**	**0.1%**	*0.0%*	*0.0%*	*0.0%*	0.0%	*0.0%*	*0.0%*
CAP Subsidy	-0.015	-0.020	-0.023	**-0.085**	**-0.089**	**-0.099**	-0.032	-0.033	-0.038
EUR 1000 change	0.2%	0.2%	0.2%	**-0.3%**	**-0.6%**	**-0.8%**	-1.0%	-0.9%	-0.9%
LFA Subsidy	0.061	0.051	0.054	**0.113**	**0.101**	**0.107**	*0.082*	0.069	*0.074*
EUR 1000 change	0.7%	0.3%	0.2%	**0.3%**	**0.4%**	**0.5%**	*2.4%*	2.0%	*2.1%*
Municipal Practice 3	**2.937**	**2.989**	**3.005**	**-1.672**	**-1.621**	**-1.543**	**4.415**	**4.378**	**4.335**
1% change	**0.4%**	**0.3%**	**0.4%**	**-0.3%**	**-0.5%**	**-0.5%**	**1.5%**	**1.7%**	**1.7%**
Municipal Practice 2	**-3.884**	**-3.960**	**-3.949**	**1.801**	**1.660**	**1.833**	**3.314**	**3.043**	**3.091**
1% change	**-2.3%**	**-2.1%**	**-2.1%**	**0.1%**	**0.2%**	**0.2%**	**2.2%**	**2.2%**	**2.1%**
Municipal No Adoption	**-8.907**	**-8.935**	**-9.093**	**-7.303**	**-7.006**	**-6.958**	**-2.636**	**-2.711**	**-2.703**
1% change	**-3.0%**	**-2.6%**	**-2.6%**	**-0.1%**	**-0.2%**	**-0.2%**	**0.6%**	**0.6%**	**0.7%**
u_{ij}	**-0.550**	-0.142	*0.578*	**-0.550**	-0.142	*0.578*	**-0.550**	-0.142	*0.578*
1 log t change	**-11.0%**	-2.1%	*8.5%*	**-0.4%**	-0.2%	*0.9%*	**-13.7%**	-3.8%	*14.8%*

Bold: significant at 5%. *Italics:* significant at 10%.

Table 11: Production estimates under MAD4 and MAD3.

	Practice 3			Practice 2			Practice 2 & 3			Not Adopting		
	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3
Constant	**14.710**	**8.507**	**8.593**	**9.712**	**9.247**	**6.839**	**9.442**	**7.766**	**10.229**	-4.771	**-2.629**	-1.292
Land	**0.880**	**0.840**	**0.809**	**0.874**	**0.850**	**0.863**	**0.907**	**0.868**	**0.896**	**0.809**	0.880*	**0.889**
Fertilizer	**0.084**	**0.100**	**0.092**	**0.077**	**0.089**	**0.084**	**0.080**	**0.085**	**0.079**	*0.055*	**0.080**	**0.072**
Machinery	-0.009	-0.005	-0.003	*0.015*	**0.016**	0.009	0.008	0.021	0.018	-0.007	*0.012*	0.018
Fuel	**0.037**	*0.030*	**0.055**	0.011	0.012	0.011	0.025	*0.033*	0.018	0.018	-0.006	0.003
2008	**-0.188**	-0.035	-0.060	**-0.151**	**-0.129**	-0.080	-0.076	-0.029	-0.096	0.166	0.126*	0.117
2009	-0.004	0.016	0.007	0.008	0.007	0.002	-0.051	0.049	0.030	0.015	*0.075*	0.075
2010	-0.079	**-0.151**	**-0.124**	0.069	0.035	-0.054	-0.053	**-0.018**	-0.008	-0.203	**-0.155**	*-0.153*
2011	0.097	0.009	0.038	0.079	*0.090*	0.024	*0.191*	0.129	0.095	-0.061	-0.049	-0.072
2012	**-0.179**	**-0.134**	**-0.143**	**-0.111**	-0.064	-0.043	0.064	**0.056**	-0.012	0.179	0.081	0.084
2013	-0.019	**-0.182**	**-0.232**	**-0.179**	**-0.241**	**-0.322**	**-0.355**	**-0.306**	**-0.275**	**-0.756**	**-0.712**	**-0.555**
Region B	0.002	0.004	**0.063**	0.038	**0.089**	**0.135**	*0.141*	0.208	**0.211**	**0.478**	0.195*	*0.192*
Region C2	-0.103	-0.035	0.001	**-0.215**	**-0.168**	**-0.101**	*-0.172*	-0.095	-0.075	**0.492**	0.201*	0.137
Region C3	**-0.332**	**-0.218**	-0.106	**-0.279**	**-0.234**	**-0.167**	*-0.215*	**-0.130**	-0.131	**0.474**	0.137*	0.113
Temperature	**-3.083**	**-1.820**	**-1.808**	**-2.008**	**-1.739**	**-1.095**	**-1.720**	**-1.488**	**-2.246**	1.621	**0.991**	0.591
Precipitation	-0.099	-0.047	-0.068	0.050	-0.030	-0.021	**-0.301**	-0.109	-0.062	0.041	-0.047	0.036
Growing Season	**-1.067**	-0.454	*-0.470*	**-0.650**	-0.687	**-0.549**	**-0.736**	-0.559	**-0.611**	0.277	0.235*	0.147
Baltic	0.040	0.016	**0.054**	**0.077**	**0.119**	**0.104**	**0.127**	0.138	**0.103**	**0.221**	*0.156*	**0.140**
σ_v	**0.310**	**0.248**	**0.190**	**0.206**	**0.225**	0.193	**0.308**	**0.308**	**0.203**	**0.450**	**0.151**	**0.169**
σ_u	**2.013**	**0.688**	**0.614**	**2.587**	**0.634**	**0.530**	**2.458**	**0.660**	**0.501**	**1.112**	0.603*	**0.453**

Bold: significant at 5%. *Italics:* significant at 10%.
*Cholesky decomposition of the inverse Hessian is numerically unstable.

Table 12: Mean inefficiency $E(u_i)$ in log t and as a percentage of full-sample mean output.

	(iii)		MAD4		MAD3	
Practice 3	1.606	3%	0.549	1%	0.490	1%
Practice 2	2.064	10%	0.506	2%	0.423	2%
Practice 2 & 3	1.961	5%	0.427	1%	0.400	1%
Not Adopting	0.887	4%	0.481	3%	0.361	2%

Table 13: $\Sigma_u^{1/2}$ estimates under the full sample (top), MAD4 (middle), and MAD3 (bottom). Correlation coefficients in parentheses.

	u_{i1}	u_{i2}	u_{i3}	u_{i4}
u_{i1}	**2.013**			
	0.688			
	0.615			
u_{i2}	**-0.585 (-0.037)**	**2.587**		
	0.097 (0.074)	**0.634**		
	0.034 (-0.050)	**0.530**		
u_{i3}	-0.544 (-0.104)	-1.704 (-0.328)	**2.458**	
	0.089 (0.059)	-0.130 (-0.161)	**0.660**	
	0.371 (0.507)	**-0.102 (-0.189)**	**0.501**	
u_{i4}	-0.434 (-0.196)	-0.752 (-0.236)	**-1.541 (-0.531)**	**1.112**
	0.177* (0.246)	-0.086* (-0.131)	**-0.166 (-0.244)**	0.603*
	0.087 (0.120)	0.001 (0.003)	-0.023 (-0.045)	**0.453**

Bold: significant at 5%. *Italics:* significant at 10%.
*Cholesky decomposition of the inverse Hessian is numerically unstable.

Table 13 presents new estimates of the correlation matrix $\Sigma_u^{1/2}$. Like the diagonal variances, the off-diagonal covariances also shrink. Most of them retain the same sign, although the correlation between practice 3 and non-adoption uniformly switches sign. Importantly, even under MAD3, some of these correlations remain large and statistically significant. This indicates that the sources of inefficiency are not themselves outlier artifacts.

This second analysis refines our initial assessment of Finland's AEP policy. As that policy was probably envisioned, the average efficient farm should not voluntarily sacrifice its efficiency by enrolling in yield-reducing AEPs. That design standard has been met according to the MAD3 analysis, but not the full-cohort analysis. The discrepancy rests with a group of exceedingly inefficient farms. Although that group comprises only about 10% of the sample, their actions are odd enough to make the whole grain sector appear more inefficient, and to send conflicting signals about policy efficacy.

8 Discussion

After correcting for adoption endogeneity, we find radically different productivity consequences for AEPs in the Finnish grain sector. Compared to standard SFA, average inefficiencies are more than twice as large with the correction in place. This rather sobering result has immediate policy implications. Indeed, many decentralized regulations besides AEPs entail voluntary choices from a menu of options, and failing to account for the adoption step can potentially overstate the welfare benefits of those policies, too.

We also find that outliers can significantly distort efficiency estimates. The outliers in our farm sample actually hold enough sway to blur our conclusions about policy efficacy. Eliminating them certainly improves the welfare picture, but it is not clear whether the trimmed cohort represents the sector as accurately. These complications are probably rather extreme, but they underscore the importance of checking for outliers as a matter of course during efficiency analysis.

Our results also inform AEP design guidelines. To the extent that efficient farms are indeed more likely to enroll in yield-reducing AEPs, the efficiency losses from those policies may be higher than originally envisioned. Consequently, more funding may be needed to compensate the increased costs that farms will actually face. Also, managers of less-efficient farms may benefit from better training on implementation, as a way to lower the barriers to adoption. With regard to Finland specifically, our results can rationalize the near-universal uptake of basic measures, but more limited use of additional and special measures. Subsidies for the latter probably do not cover a typical farm's potential decline in productivity.

References

Dennis Aigner, C.A. Knox Lovell, and Peter Schmidt. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1):21–37, 1977.

Christine Amsler, Artem Prokhorov, and Peter Schmidt. Endogeneity in stochastic frontier models. *Journal of Econometrics*, 190(2):280–288, 2016.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Laure Bamière, Petr Havlík, Florence Jacquet, Michel Lherm, Guy Millet, and Vincent Bretagnolle. Farming system modelling for agri-environmental policy design: The case of a spatially non-aggregated allocation of conservation measures. *Ecological Economics*, 70(5):891–899, 2011.

Kathy Baylis, Stephen Peplow, Gordon Rausser, and Leo Simon. Agri-environmental policies in the EU and United States: A comparison. *Ecological Economics*, 65(4): 753–764, 2008.

Sylvain Chabé-Ferret and Julie Subervie. How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. *Journal of Environmental Economics and Management*, 65(1):12–27, 2013.

European Commission. Agri-environment measures: Overview on general principles, types of measures, and application. Directorate General for Agriculture and Rural Development, 2005.

Hongli Feng, Lyubov A. Kurkalova, Catherine L. Kling, and Philip W. Gassman. Environmental conservation in agriculture: Land retirement vs. changing practices on working land. *Journal of Environmental Economics and Management*, 52(2):600–614, 2006.

Thomas Gerstner and Michael Griebel. Numerical integration using sparse grids. *Numerical Algorithms*, 18(3):209, 1998.

John Geweke. Monte Carlo simulation and numerical integration. volume 1 of *Handbook of Computational Economics*, pages 731 – 800. Elsevier, 1996.

Eirini Giovanopoulou, Stefanos A. Nastis, and Evagelos Papanagiotou. Modeling farmer participation in agri-environmental nitrate pollution reducing schemes. *Ecological Economics*, 70(11):2175–2180, 2011.

Frank R. Hampel. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393, 1974.

James J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1): 153–161, 1979.

William C. Horrace, Xiaodong Liu, and Eleonora Patacchini. Endogenous network production functions with selectivity. *Journal of Econometrics*, 190(2):222–232, 2016.

Stephen Hynes, Niall Farrelly, Eithne Murphy, and Cathal O'Donoghue. Modelling habitat conservation and participation in agri-environmental schemes: A spatial microsimulation approach. *Ecological Economics*, 66(2–3):258–269, 2008.

Michael P. Keane. A computationally practical simulation estimator for panel data. *Econometrica*, 62(1):95–116, 1994.

Subal C. Kumbhakar and C.A. Knox Lovell. *Stochastic Frontier Analysis.* Cambridge University Press, 2000.

Subal C. Kumbhakar, Efthymios G. Tsionas, and Timo Sipiläinen. Joint estimation of technology choice and technical efficiency: An application to organic and conventional dairy farming. *Journal of Productivity Analysis*, 31(3):151–161, 2009.

Marita Laukkanen and Céline Nauges. Evaluating greening farm policies: A structural model for assessing agri-environmental subsidies. *Land Economics*, 90(3):458–481, 2014.

Sébastien Mary. Assessing the impacts of Pillar 1 and 2 subsidies on TFP in French crop farms. *Journal of Agricultural Economics*, 64(1):133–144, 2013.

Rosa L. Matzkin. Nonparametric identification. volume 6 of *Handbook of Econometrics*, pages 5307–5368. Elsevier, 2007.

Cornelia Ohl, Martin Drechsler, Karin Johst, and Frank Wätzold. Compensation payments for habitat heterogeneity: Existence, efficiency, and fairness considerations. *Ecological Economics*, 67(2):162–174, 2008.

Isabelle Piot-Lepetit and Monique Le Moing. Productivity and environmental regulation: the effect of the nitrates directive in the French pig sector. *Environmental and Resource Economics*, 38(4):433–446, 2007.

Andrea Pufahl and Christoph R. Weiss. Evaluating the effects of farm programmes: results from propensity score matching. *European Review of Agricultural Economics*, 36(1):79–101, 2009.

Sergey Rabotyagov, Todd Campbell, Manoj Jha, Philip W. Gassman, Jeffrey Arnold, Lyubov Kurkalova, Silvia Secchi, Hongli Feng, and Catherine L. Kling. Least-cost control of agricultural nutrient contributions to the Gulf of Mexico hypoxic zone. *Ecological Applications*, 20(6):1542–1555, 2010.

Wolfram Schlenker and Michael J. Roberts. Estimating the impact of climate change on crop yields: the importance of nonlinear temperature effects. Working Paper 13799, National Bureau of Economic Research, 2008.

James H. Stock, Jonathan H. Wright, and Motohiro Yogo. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529, 2002.

Lota D. Tamini, Bruno Larue, and Gale West. Technical and environmental efficiencies and best management practices in agriculture. *Applied Economics*, 44(13):1659–1672, 2012.

Frank Wätzold and Martin Drechsler. Spatially uniform versus spatially heterogeneous compensation payments for biodiversity-enhancing land-use measures. *Environmental and Resource Economics*, 31(1):73–93, 2005.

Frank Wätzold, Nele Lienhoop, Martin Drechsler, and Josef Settele. Estimating optimal conservation in the context of agri-environmental schemes. *Ecological Economics*, 68 (1-2):295–305, 2008.

Jinhua Zhao, Lyubov A. Kurkalova, and Catherine L. Kling. Alternative green payment policies when multiple benefits matter. *Agricultural and Resource Economics Review*, 33(1):148–158, 2004.

Appendix A Numerical Integration

Numerical integration of high-dimensional integrals suffers from rather severe tractability issues (Geweke, 1996). We discuss two integration algorithms below, one that is familiar to econometricians and one that is perhaps less well-known. They exhibit very different degrees of approximation error in our model.

The first is an integration-by-simulation algorithm based on the Geweke-Hajivassiliou-Keane (GHK) simulator (Keane, 1994). This involves sampling from an easier, related distribution $g(u_i)$ – the multivariate normal in our case – and mapping those samples to $f(u_i)$. Letting $M(u_i)$ denote the integrand of the likelihood, the relationship that enables the simulator is the identity

$$\int_{\mathbb{R}^{J+}} M(u_i)\, dF(u_i) = \int_{\mathbb{R}^{J}} M(u_i) \frac{f(u_i)}{g(u_i)} g(u_i)\, du_i$$

By computing each simulation's importance weight $w(u_i) = \frac{f(u_i)}{g(u_i)}$, the integral can be approximated with S simulations from $g(u_i)$:

$$\int_{\mathbb{R}^{J+}} M(u_i)\, dF(u_i) \simeq \frac{\frac{1}{S}\sum_{s=1}^{S} M(u_i^s)\, w(u_i^s)}{\frac{1}{S}\sum_{s=1}^{S} w(u_i^s)}$$

The numerator approximates the likelihood for an untruncated, joint-normal distribution, and the denominator approximates the truncation probability.

Though easy to implement, GHK exhibits unacceptable approximation error in our model. In explorations using the toy calibration from Section 4, we find that obtaining two digits of integration accuracy requires $S \approx 10^4$ points. Because integration-by-simulation relies on convergence in probability, any further improvements come at a $\sqrt{S}$ rate at best. This implies that an exponentially-larger number of simulations would be needed to achieve each additional digit of accuracy, a quintessential curse-of-dimensionality result. The finding bodes particularly ill for our application: the dataset has $I = 10^3$ to 10^4 observations, each of which entails an integration at every iteration of a maximum-likelihood search.

Instead, we perform the integration using a sparse grid, a set of multidimensional quadrature points and associated scalar weights $\{\omega_i^s, q_i^s\}_{s=1}^S$. Sparse grids are constructed under a different convergence criterion that reduces integration error much faster (Gerstner and Griebel, 1998). As a result, a smaller number of points S is needed to achieve the same level of precision. The functional form of the approximation is like a weighted average:

$$\int_{\mathbb{R}^{J+}} M(u_i)\, dF(u_i) \simeq \sum_{s=1}^{S} M(\omega_i^s)\, f(\omega_i^s)\, q_i^s$$

In experiments with the toy calibration, we obtain 3 to 4 digits of accuracy with a grid of $S \approx 10^2$ points, a far more tractable option.

These grids are defined over simple regions of the real domain, and the quadrature points must be analytically mapped to more complicated regions of interest like ours. We start from the Gauss-Legendre grid, which approximates integration over the $[-1,1]^J$ hypercube, and map its points to $\mathbb{R}^{J+}$ while imputing the correct correlation structure. To frame the procedure in better-known terms, this translation treats each point ω_i^s in the original grid as a "draw from a uniform distribution" $\bar{\omega}_i^s$, to which the "inverse distribution function" $F^{-1}(\bar{\omega}_i^s)$ is applied. The result is analogous to a "draw from the target distribution" u_i^s.

The procedure involves two major domain transformations:

1. The original grid point ω_i^s in $[-1,1]^J$ is linearly translated to a point $\bar{\omega}_i^s$ in the uniform-distribution domain $[0,1]^J$.

2. The uniform point $\bar{\omega}_i^s$ is translated to $F(u_i^s)$. Because this distribution has arbitrary correlation as well as truncation, the u_i^s vector must be composed sequentially, with each element accounting for the correlation previously imposed.

    (a) For the jth element in the vector, an intermediate scalar draw $\bar{u}_{ij}^s$ is taken from a standard univariate truncated-normal distribution, using the uniform scalar $\bar{\omega}_{ij}^s$ from step 1. To capture the correlation in elements up to the jth, the lower truncation bound of $\bar{u}_{ij}^s$ is set to the z-score $\frac{0-\sum_{j'=1}^{j-1}\sigma_{ujj'}\bar{u}_{ij'}^s}{\sigma_{ujj}}$, where $\sigma_{ujj'}$ denotes the (j,j') element of the correlation matrix $\Sigma_u^{1/2}$.

(b) Repeating (a) for all J elements results in a correlated, truncated standard draw $\bar{u}_i^s$. This is translated into a draw from $N^+(0, \Sigma_u)$ through the affine transformation $u_i^s = 0 + \sum_u^{1/2} \bar{u}_i^s$.

Very importantly, the change-of-variable formula for integration must be applied in parallel with every transformation of the grid point. This ensures that the Gauss-Legendre integrand $f(\omega_i)\, d\omega_i$ over $[-1, 1]^J$ corresponds to the multivariate truncated-normal integrand $f(u_i)\, du_i$ over $\mathbb{R}^{J+}$.

Appendix B First-Stage Estimates

Table 14: Multinomial logit (a) and endogeneity-corrected (iii) adoption estimates (continues Table 8).

	Practice 3		Practice 2		Practice 2 & 3	
	(a)	(iii)	(a)	(iii)	(a)	(iii)
Manager's Age	0.001	0.002	-0.001	0.000	-0.004	-0.004
Debts/Assets	**0.244**	**0.243**	*0.089*	*0.089*	*0.098*	0.096
CAP Subsidy	-0.017	-0.015	**-0.086**	**-0.085**	-0.032	-0.032
LFA Subsidy	0.062	0.061	**0.113**	**0.113**	*0.082*	*0.082*
Municipal Practice 3	**2.934**	**2.937**	**-1.632**	**-1.672**	**4.399**	**4.415**
Municipal Practice 2	**-3.789**	**-3.884**	**1.758**	**1.801**	**3.243**	**3.314**
Municipal No Adoption	**-8.752**	**-8.907**	**-6.939**	**-7.303**	**-2.578**	**-2.636**
u_{ij}		**-0.550**		**-0.550**		**-0.550**
Constant	11.228	6.047	-2.291	-6.735	5.893	2.006
Land	**1.081**	**1.088**	0.276	0.281	**0.591**	**0.598**
Fertilizer	0.006	0.006	-0.038	-0.040	0.131	0.132
Machinery	**0.221**	**0.223**	0.069	0.068	**0.193**	**0.192**
Fuel	0.113	0.112	0.086	0.085	**0.189**	**0.190**
2008	-0.231	-0.114	0.017	0.130	-0.257	-0.169
2009	-0.097	-0.083	-0.045	-0.041	-0.027	-0.019
2010	0.533	0.454	0.071	0.014	0.515	0.437
2011	*0.644*	0.529	-0.120	-0.198	*0.671*	0.579
2012	*0.594*	*0.631*	0.431	0.456	0.531	*0.566*
2013	0.323	0.267	-0.480	-0.541	0.013	-0.069
Subsidy region 2	*-0.746*	*-0.707*	**-1.584**	**-1.557**	**-1.016**	**-0.960**
Subsidy region 3	-0.467	-0.362	**-1.762**	**-1.696**	**-1.787**	**-1.699**
Subsidy region 4	**-1.424**	**-1.292**	**-1.926**	**-1.828**	**-1.991**	**-1.897**
Temperature	-2.149	-0.995	2.226	3.148	0.074	0.947
Precipitation	-0.777	-0.745	-0.225	-0.198	-0.801	-0.781
Growing Season	-1.629	-1.215	-0.284	0.035	*-2.209*	-1.902
Baltic	0.141	0.160	0.073	0.069	0.169	0.182

Bold: significant at 5%. *Italics:* significant at 10%.

Table 15: Selection-equation point estimates under MAD4 and MAD3 (continues Table 10).

	Practice 3			Practice 2			Practice 2 & 3		
	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3	(iii)	MAD4	MAD3
Manager's Age	0.002	0.000	0.001	0.000	-0.002	-0.001	-0.004	-0.004	-0.003
Debts/Assets	**0.243**	**0.253**	**0.257**	*0.089*	*0.102*	*0.102*	0.096	*0.102*	*0.114*
CAP Subsidy	-0.015	-0.020	-0.023	**-0.085**	**-0.089**	**-0.099**	-0.032	-0.033	-0.038
LFA Subsidy	0.061	0.051	0.054	**0.113**	**0.101**	**0.107**	*0.082*	0.069	*0.074*
Municipal Practice 3	**2.937**	**2.989**	**3.005**	**-1.672**	**-1.621**	**-1.543**	**4.415**	**4.378**	**4.335**
Municipal Practice 2	**-3.884**	**-3.960**	**-3.949**	**1.801**	**1.660**	**1.833**	**3.314**	**3.043**	**3.091**
Municipal No Adoption	**-8.907**	**-8.935**	**-9.093**	**-7.303**	**-7.006**	**-6.958**	**-2.636**	**-2.711**	**-2.703**
u_{ij}	**-0.550**	-0.142	*0.578*	**-0.550**	-0.142	*0.578*	**-0.550**	-0.142	*0.578*
Constant	6.047	6.043	7.728	-6.735	-7.158	-7.088	2.006	2.344	0.831
Land	**1.088**	**1.184**	**1.117**	0.281	*0.388*	0.357	**0.598**	**0.719**	**0.673**
Fertilizer	0.006	0.060	0.065	-0.040	-0.001	0.011	0.132	*0.164*	*0.172*
Machinery	**0.223**	**0.227**	**0.218**	0.068	*0.076*	0.073	**0.192**	**0.192**	**0.203**
Fuel	0.112	0.091	0.105	0.085	0.071	0.072	**0.190**	*0.161*	*0.161*
2008	-0.114	-0.070	-0.106	0.130	0.144	0.126	-0.169	-0.149	-0.150
2009	-0.083	-0.106	-0.115	-0.041	-0.051	0.037	-0.019	-0.112	-0.047
2010	0.454	0.489	0.512	0.014	0.052	0.075	0.437	0.492	0.435
2011	0.529	0.520	0.537	-0.198	-0.241	-0.239	0.579	0.596	0.565
2012	*0.631*	**0.692**	*0.608*	0.456	0.462	0.465	*0.566*	*0.581*	0.506
2013	0.267	0.226	0.252	-0.541	-0.536	-0.477	-0.069	-0.108	-0.119
Region B	*-0.707*	**-0.924**	**-0.909**	**-1.557**	**-1.789**	**-1.706**	**-0.960**	**-1.204**	**-1.128**
Region C1	-0.362	-0.475	-0.499	**-1.696**	**-1.848**	**-1.888**	**-1.699**	**-1.878**	**-1.898**
Region C2	**-1.292**	**-1.431**	**-1.460**	**-1.828**	**-1.996**	**-2.061**	**-1.897**	**-2.024**	**-2.044**
Temperature	-0.995	-0.801	-1.241	3.148	3.188	3.210	0.947	0.989	1.268
Precipitation	-0.745	-0.753	-0.674	-0.198	-0.149	-0.041	-0.781	-0.879	-0.658
Growing Season	-1.215	-1.330	-1.439	0.035	0.142	0.086	-1.902	-1.983	-1.893
Baltic	0.160	0.124	0.106	0.069	0.070	0.106	0.182	0.162	0.204

Bold: significant at 5%. *Italics:* significant at 10%.